

**The California Central Coast Research Partnership:  
Building Relationships, Partnerships and Paradigms for University-Industry  
Research Collaboration.**

FINAL REPORT ON ONR GRANT NO. N00014-05-1-0855

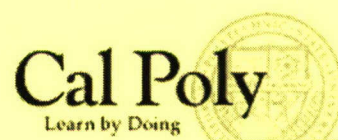
PERIOD OF PERFORMANCE: June 15, 2005 through January 31, 2007

Principal Investigator:

Susan Opava, Ph.D.  
Dean of Research and Graduate Programs  
California Polytechnic State University  
San Luis Obispo, California



CALIFORNIA CENTRAL COAST  
RESEARCH PARTNERSHIP



**The California Central Coast Research Partnership:  
Building Relationships, Partnerships and Paradigms for University-Industry  
Research Collaboration.**

FINAL REPORT ON ONR GRANT NO. N00014-05-1-0855

PERIOD OF PERFORMANCE: June 15, 2005 through January 31, 2007

Principal Investigator:

Susan Opava, Ph.D.  
Dean of Research and Graduate Programs  
California Polytechnic State University  
San Luis Obispo, California

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 26/04/2007		2. REPORT TYPE Final Technical Report		3. DATES COVERED (From - To) 6/15/2005 - 1/31/2007	
4. TITLE AND SUBTITLE The California Central Coast Research Partnership: Building Relationships, Partnerships and Paradigms for University-Industry Research Collaboration				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER N00014-05-1-0855	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Opava, Susan C.; Adams, Nikki; Arakaki, Dean; Beckett, Jonathon; Bensky, Thomas; Carpenter, Thomas; Chadwell, Charles; Crockett, Robert; Davol, Andrew; Echols, Robert; Fernando, Raymond; Fiegel, Gregg; Franklin, Diana; Freed, Tali; Gollery, Steven; Hanson, James; Harfenist, Steven; Immoos, Chad; Jimenez-Flores, Rafael; Jin, Xiaomin; Klisch, Steven; Liddicoat, Albert; Lindert, Lisa; Marshall, David; (Continued on reverse side)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Cal Poly Corporation Sponsored Programs Dept. Bldg. 38 Room 102 San Luis Obispo, CA 93407-0830				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Clifford W. Anderson Office of Naval Research 875 N. Randolph St. One Liberty Center Arlington, VA 22203-1995				10. SPONSOR/MONITOR'S ACRONYM(S) ONR	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) 353	
12. DISTRIBUTION/AVAILABILITY STATEMENT Available for public release					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The primary purpose of this initiative is to forge a strong link between private sector R&D and University applied research, to speed the development of new knowledge and the transfer of technology to the public and private sectors. To this end, communications infrastructure and R&D facilities have been developed. Relationships with private companies engaged in R&D have been advanced. Important research has been carried out in areas of interest to the DOD and national security. These areas include: Nanotechnology, hyperspectral imaging and analysis, smart materials, energy efficiency, communications, network technologies, sensors, malaria, ultraviolet irradiation, biohazards, chemical agents, RFID, visual speech analysis, artificial neural networks, fluid dynamics, and underwater optical sensing.					
15. SUBJECT TERMS Nanotechnology, hyperspectral imaging and analysis, smart materials, energy efficiency, communications, network technologies, sensors, malaria, ultraviolet irradiation, biohazards, chemical agents, RFID, visual speech analysis, artificial neural networks, fluid dynamics, and underwater optical sensing.					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT  UU	18. NUMBER OF PAGES  655	19a. NAME OF RESPONSIBLE PERSON Susan C. Opava, Ph.D.
a. REPORT U	b. ABSTRACT U	c. THIS PAGE UU			19b. TELEPHONE NUMBER (Include area code) 805-756-1508

Answer 6, continued:

Meagher, James; Mehiel, Eric; Merayyan, Saad; Moline, Mark; Moss, Robb; Murray, William; Pan, Jianbiao; Peterson, Dan; Pohl, Jens; Rahim, Ashraf; Saunders, Karl; Savage, Richard; Schwartz, Peter; Seng, John; Sharpe, John; Singh, Jay; Waldorf, Daniel; Wu, Xi; Yu, Xiao-Hua; Zhang, Jane

## TABLE OF CONTENTS

I.	Title of Project and Principal Investigator .....	3
II.	Summary of Project .....	3
III.	Relevance to ONR Objectives .....	3-6
	A. Relevant partners .....	3-4
	B. Relevant R&D focus .....	5
	C. University-industry-government partnership .....	5
	D. University strengths .....	6
IV.	Summary of Results During the Period of Performance .....	6-655
	A. General .....	6-9
	B. Development of new research capacity.....	9-27
	C. Detailed research reports .....	28-655
	1. Results of CADRC research	
	• Knowledge management using semantic web languages and technologies..	29-81
	2. Results of research projects .....	82-655
	• Molecular sensors and defenses against ultraviolet radiation.....	82-121
	• Development of wideband microstrip geometries, RFID tag mounting structures and radiated emissions testing capabilities.....	122-135
	• Validation of RFID livestock identification systems.....	136-146
	• Continuous optical monitoring of a near-shore sea-water column.....	147-155
	• Seismic investigation of the rigid body rocking of portable steel wine barrel stacks.....	156-201
	• Three-dimensional scaffolds for mammary epithelial cell growth.....	202
	• Improving power conversion efficiencies of polymer based solar cells.....	203-212
	• Development of compaction procedures for optimum placement of wastes in municipal solid waste landfills using GPS.....	213-230
	• Production, characterization and self-assembly of passivated and functionalized metal nanoparticles and collaborative use of Zeiss-10 transmission electron microscope.....	231-236
	• Development of polymer coatings for chemical weapons degradation.....	237-241
	• VCSEL-based optical transmitter system using injection-locking tech.....	242-259
	• Nonlinear constitutive equations for cartilage growth analysis.....	260-280
	• Delay and disruption tolerant network implementation with reconfigurable hardware support.....	281-292
	• Mechanistic study of the cyclin-dependent kinases PfPK5/Pfcyc-1 and PfPK6 from the malarial parasite, <i>Plasmodia falciparum</i> .....	293-297
	• Development of a parallel computational framework for computational fluid dynamic simulations.....	298-310
	• The Cal Poly spacecraft attitude dynamics simulator.....	311-401
	• Assessment of contaminant transport into groundwater from selected landfill sites in the state of California.....	402-480
	• Evaluating the effects of aging on liquefiable soil deposits.....	481-493
	• Development of hybrid rocket motor facility for advanced nozzle research	494-502

## 2. Results of research projects – Continued

- Effect of reflow profile and thermal shock on lead-free solder reliability... 503-518
- Recycled waste materials as additives to improve the performance of soil-cement: A laboratory investigation..... 519-548
- Modeling the strong electro-optical response of "de Vries" ferroelectric liquid crystals..... 549-557
- Development of a micro system actuator employing shape memory alloys for exploring the nanotechnology world..... 558-572
- DNA-linked polystyrene microstructures: Selective DNA attachment and enhanced electrostatic stabilization..... 573-579
- Large-scale distributed wireless networks..... 580-587
- Passive directionally sensitive laser Doppler velocimeter..... 588-593
- Nano-grain cutting tool for machining high-temp Aerospace materials..... 594-603
- Crack diagnosis of a rotor from torsional/lateral coupling mechanisms..... 604-623
- Adaptive controller design with artificial neural networks..... 624-645
- Visual speech analysis in real-life environment..... 646-655

**FINAL REPORT ON ONR GRANT NO. N00014-05-1-0855**  
**PERIOD OF PERFORMANCE: 6/15/05 to 1/31/07**

**I. Title of Project and Principal Investigator**

The California Central Coast Research Partnership: Building Relationships, Partnerships and Paradigms for University-Industry Research Collaboration, Susan C. Opava, Ph.D.

**II. Summary of Project**

The mission of the California Central Coast Research Partnership (C<sup>3</sup>RP) is to facilitate the exchange of knowledge and skills between the higher education sector and the private sector in San Luis Obispo County, and to encourage the growth of high-tech companies in the region, thereby enhancing economic development and quality of life. The partnership is a long-term plan to create a dynamic and self-supporting university-industry-government partnership that capitalizes on the strengths and mutual interests of the educational and technology-based business sectors. The plan recognizes the key role of higher education in preparing a highly skilled work force and transferring new knowledge to practical uses. The outcomes of this partnership, when fully realized, will be the expansion of current and the creation of new University technological R&D activities; the development of existing technology-based businesses and the creation of new ones; and the generation of opportunities for job training and research and development activities for University and Community College students and faculty.

The project will eventually lead to the construction (with non-DOD funding) of a technology park on the California Polytechnic State University campus, which will provide state-of-the-art space for private technology companies engaged in research and development activities, as well as a business incubator that will provide all of the support services needed by start-up, technology-based companies. The aspect of the program supported by this project, is the continued development of a strong base of applied research at Cal Poly, through university-government-industry partnerships designed to optimize the application of the strengths of each of these sectors to problems of mutual interest. The management team, operational since January '02, continues to lead the project and develop the collaborative relationships between the educational and private sectors that are essential to realizing long-term goals and securing the financial base that will allow full-scale project development.

**III. Relevance to ONR Objectives**

**A. Relevant partners.**

C<sup>3</sup>RP represents a coalition of educational institutions, local, state and federal government, and private businesses that have worked together in unprecedented fashion to advance the common goals inherent in the proposed university-industry partnership. The current partners in the project and their contributions include:

**California Polytechnic State University**

- committed the land for the project, valued at ~\$1.5 million
- provided assistance in financial management of the project
- contributed \$90,000 for a pre-feasibility study by Bechtel Corporation
- committed several hundred thousand dollars of in-kind contributions of senior management time and effort over several years; continues to do so
- invested ~\$1,000,000 in efforts to raise additional funds for the project

**GEO, International** (GEOgraphic Network Affiliates, International; a private company)

- has worked *pro bono* with C<sup>3</sup>RP on communications aspects of the project

**CENIC** (Corporation for Educational Network Initiatives in California; association of Internet2 universities in CA)

- works with Cal Poly to further the goals of the IEEAF (see below), goals that will directly benefit C<sup>3</sup>RP

**IEEAF** (International Educational Equal Access Foundation; established by GEO and CENIC)

- has secured donations of virtual and physical communications assets in at least 37 countries; some of these fiber-optic assets directly benefit this ONR project

**City of San Luis Obispo**

- in partnership with Cal Poly has developed a carrier-neutral fiberoptic ring around the city.

**National Science Foundation**

- is working with GEO, IEEAF, CENIC and other universities in the United States to promote the goals of these organizations to develop low-cost fiber-optic networks for the benefit of educational institutions, non-profit organizations and local communities.

**Housing and Urban Development**

- Has provided funds toward construction of the pilot technology park building.

**Economic Development Administration**

- Has provided funds toward construction of the pilot technology park building.

Efforts are ongoing to secure new partners, including:

- Major corporations
- Small technology-based businesses

## **B. Relevant R&D focus.**

The research programs that were supported are relevant to seven of the eight “thrust areas” of ONR’s Code 30 Science and Technology Program. The projects involved basic research in these areas, as well as applied research and development leading to more immediate technological applications. The seven areas of relevance and the more specific focus areas to which the research contributed are listed below:

**Command and Control, Computers, Communication:** information security; situational awareness; disruption and delay tolerant networks; robust communications; conformal antennae; GPS applications; visual speech analysis; large-scale, distributed wireless networks.

**Force Protection:** improved protection for the individual; smart materials; injury mitigation; IED detection; bio- and chemical-hazard detection and mitigation.

**Mine Countermeasures:** shallow underwater imaging; hyperspectral imaging; IED detection; situational awareness.

**Human Performance, Training and Survivability:** cognitive performance enhancement; physical performance enhancement; smart materials; sensors; natural sunscreens and irradiation markers.

**Intelligence, Surveillance and Reconnaissance:** sensor fields; data/information analysis and distribution; tagging, tracking and locating; antennae.

**Logistics:** RFID for logistics control; alternative power sources; solar cells.

**Maneuver:** advanced and composite materials for vehicles and aircraft; new materials and designs for survivability.

## **C. University-industry-government partnership.**

The primary focus of this initiative is to forge a strong link between private sector R&D and University applied research to speed the development of new knowledge and the transfer of technology to the public and private sectors. San Luis Obispo has become a draw for technology businesses (with a heavy concentration of software development companies) from both the LA Basin and Silicon Valley. For example, SRI (Stanford Research Institute), International operates a "software center of excellence" in the city. Branches of major corporations are also located nearby, for example, Sun Microsystems, Veritas, and Sunbay Software. Lockheed-Martin has a research and development group in nearby Santa Maria. Two local companies manufactured critical components for the Mars rovers and other companies, e.g. California Fine Wire, are suppliers to the military. Also located on the Central Coast are branches of two major biotechnology companies: Promega Biosciences and Santa Cruz Biotechnology.

#### **D. University strengths.**

Cal Poly is a State university that has achieved national distinction as a polytechnic university, with engineering and computer science programs ranked among the very best undergraduate programs in the country. Its strengths have led it to orchestrate the research partnership effort and the consortium of partners described herein. Cal Poly also has affiliations with CSA (California Space Alliance) and with Vandenberg Air Force Base, where it has offered an M.S. in Aerospace Engineering by distance learning. The high bandwidth that will be associated with the eventual physical site selected for the partnership will allow Cal Poly to offer many more academic programs by distance learning to remote locations. In particular, through possible collaborative agreements at cable-head locations around the world (including Asia and Europe) our programs can be made available to military personnel stationed almost anywhere in the world. This could be tied into training programs for ONR and, if desirable, to training directly related to research projects.

**In summary**, the California Central Coast Research Partnership has taken advantage of a confluence of factors, including existing and potential relationships, fortuitous and unique technological and economic developments in the region, the particular strengths and expertise of the CAD Research Center, and a meshing of the research and development interests of the University, the Office of Naval Research, and the private sector. C<sup>3</sup>RP is the vehicle for fully realizing the benefits of the common goals and synergies of the partners and their respective resources.

### **IV. Summary of Results During the Period of Performance**

#### **A. General.**

This program was originally funded through an award from ONR in FY '02, covering the period September '01 through December '02. Accomplishments during this first award period were described in a detailed report to ONR in March '03. A second award began in June '02 and ended in December '03. Accomplishments during this second award period were described in a detailed report to ONR in March '04. A third award began in March '04 and ended in July '05. Accomplishments during this third award period were described in a detailed report to ONR in October '05. This report covers a fourth award period, which began on 6/15/2005 and ended on 1/31/2007. General accomplishments are summarized below. Detailed reports are presented later in the document.

An overview of accomplishments during this project period follows:

- Research carried out by the **CADRC (Cooperative Agent Design Research Center)**, of particular interest to ONR and the Marine Corps, was again funded. A detailed report on this project, Knowledge Management Using Semantic Web Languages and Technologies, is provided later on in this document.

- **New research** has been developed and some research has been continued, including some with industry collaboration. Projects address topics highly relevant to defense and national security, such as nanotechnology, hyperspectral imaging and analysis, smart materials, energy efficiency, communications, network technologies, sensors, malaria, ultraviolet irradiation, biohazards, chemical agents, RFID, visual speech analysis, artificial neural networks, fluid dynamics, and underwater optical sensing. Detailed reports of the results of these projects are presented in Section IV. C. of this report.
- Support for these research projects has resulted in **significant follow-on funding** from government and private sponsors. Since January 1, 2003, C3RP-supported faculty have received almost \$25 million (\$24,989,079) in competitive funding from other sources.
- Research continued on a **new technology**, developed in the previous award period, that uses fiberoptics for undersea optical sensing.
- **New research capacity was developed**, including new instrumentation and enhanced infrastructure, as well as capacity in new research areas (detailed below).
- Significant enhancements have been made in **information technology infrastructure**:
  - **Internet2** connectivity was initially applied for, approved, and acquired for the campus in November 2001, to support current and future research efforts. Internet2 membership and connectivity has continued this year
  - CENIC (see above) and AARNET, Australia's Academic and Research Network (<http://aarnet.edu.au>) continue to work on a collaborative effort with the private sector, to implement a **trans-Pacific high-speed network connection (10 Gigabits/sec)**. The California-based landing site will be in San Luis Obispo County and will connect into CENIC's high-speed network at its network point of presence located on the Cal Poly campus. Cal Poly's geographic location, its ability to act as a "hardened" facility to support critical network backbone equipment, as well as the existing fiber-optic infrastructure built out in partnership with the City of San Luis Obispo, has positioned Cal Poly to be the access point for the high-speed network corridor to Australia's research environments and potentially other Eastern Pacific Rim research networks. **This project is being developed without use of ONR funds.**
  - A **grid computing network** was established during the previous grant period, to support specific research projects that require high-end computing capability previously unavailable on campus. During this period, the IBM e-Cluster was installed and made available for faculty use. Staff were assigned to support the environment and trained on Grid services. Faculty were notified of the service and a web page describing the service put in place: <http://gridcomputing.calpoly.edu>. One faculty

member has utilized the environment for computation and four more have shown interest. Documentation of support procedures is in progress. Hardware and software evaluation of the IBM e-cluster environment was performed to determine any additional component requirements to better support the environment.

Next steps include adding other users to test management of resources in a shared environment, searching for candidates to utilize the IBM Deep Computing Center, establishing scheduling and service expectations to deal with resource contention.

- Dr. Franz Kurfess, Professor of Computer Science, participated in the Fall 2006 Internet2 Member Meeting in Chicago, IL. Professor Kurfess attended the meeting of the Program Committee; participated in the Teaching and Learning Special Interest Group; planned future 'neternship' activities and identified potential host organizations for Cal Poly 'neterns'; and met with a group of people interested in pursuing a course proposal on teaching advanced technology for general education by utilizing advanced networking infrastructure like Internet2.
- Under the supervision of Dr. Franz Kurfess, four students became the first Internet2 Neternship team at Cal Poly. The Neternship required them to collaborate virtually with an Internet2 contact in Michigan to solve a problem identified by the Internet2 community. Their task was to create a bulk file-transfer application for high performance networks that would increase data transfer rates and be easy to use for novices. For six months, the team worked closely with various contacts from Internet2, and other networking professionals around the country, to design and develop the application, leveraging Internet2 resources to remotely collaborate with others. The team successfully completed research on a backend for their application that increased data transfer rates over typical file-transfer protocols by almost 2000%, and developed a user-centered graphical user interface for the application.

Two of the students (Rachelle Hom and Gigi Choy) presented posters showcasing the work at the Spring Internet2 Member Meeting. The posters were among only fifteen approved to be shown at the meeting. One poster described the Neternship experience; the other described the technical aspects of ÜberSCP, the bulk file-transfer application for high performance networks.

- A **database of technology-based companies** that are potential partners in the project and research collaborators has been updated and expanded.
- The first **research and development company has located on campus** in anticipation of the construction of the pilot building for the technology park. Applied Biotechnology, Inc. is a company specializing in the use of genetically modified

plants to produce non-food products, for example, industrial enzymes, biochemical reagents and oral vaccines. The presence of the company has catalyzed interest of faculty in developing research in this area.

- **Relationships** have been established with new technology companies that are potential research collaborators, including: Novariant, Vacca, Inc. Innovative Energy Solutions, Agilence, Nature's Fuel, Moog, Nellcor (a division of Tyco Healthcare), Xcelaero, and Parsons.
- The **web site for the project** ([www.c3rp.org](http://www.c3rp.org)), which presents C<sup>3</sup>RP as an interface between Cal Poly and business/industry for the purpose of facilitating R&D relationships, has been maintained and updated.
- A **research brochure**, highlighting the applied research strengths, facilities and resources of Cal Poly, was completed. It reports results of ongoing representative projects and will be a tool for attracting industry partners to collaborate in the applied research efforts being supported through C<sup>3</sup>RP.
- Efforts continue to develop industry partners in the **biotechnology sector** for the purpose of developing research and training activities in this field. To this end we continue to work with the Central Coast Biotechnology Center in Ventura, CA, with two local community colleges, and with several biotech companies, including Amgen, Baxter, Fziomed, Genentech, Promega BioSciences, Hardy Diagnostics, and Santa Cruz Biotechnology.
- The project's leaders have continued to work with other private and government partners to advance the project and to attract research collaborators and support. During this project period, we were successful in obtaining an **award from the Economic Development Administration** of \$1.8 million for construction of the first building for the technology park.

## **B. Development of new research capacity**

One of the goals of the project was to increase the capacity of the organization to carry out state-of-the-art research in the areas of interest. To this end, specialized instrumentation was acquired, infrastructure was developed, and research funding was provided to seed new areas of emphasis.

### **1. Instrumentation.**

In the area of instrumentation, we acquired the following equipment/systems for shared use. Other instrumentation, acquired for use on individual projects, is described in the reports for those projects.

- Typhoon variable-mode, scanning imaging system, to support several projects with proteomics components

- Circulating seawater system (partial funding provided), to support several research projects at the Avila pier
- Atomic Force Microscope (research grade), to support nanotechnology research projects
- Lyophilizer (partial funding provided), to support biotechnology-related research projects
- Transmission Electron Microscope (purchased used), to support projects in nanotechnology and materials science
- Mass spectrometer (partial funding provided), to support a number of C<sup>3</sup>RP-funded research projects

## **2. Infrastructure.**

In the infrastructure area, the following project was supported.

### ***High Bandwidth Imagery and Data Fusion Geospatial Infrastructure, PI: Mark Moline***

#### *Background*

*Data fusion is the process by which two or more streams of data of different temporal and spatial resolution are combined to produce synthesized products which are unattainable from a discrete image. Data fusion allows for the more complete imaging of an area intermittently covered by clouds. It may also provide better analysis of movement or time-dependent change in targets and/or environmental conditions. In the electromagnetic frequency domain, targets which are visible in one image type may be less than visible in another, and vice versa. Synthetic Aperture Radar (SAR) imagery, for example, may illuminate target elevations or directional wave spectra, whereas optical imagery may identify targets by shape and color. The abilities to identify bathymetry, submerged obstacles, bottom type, or water clarity are clearly required for Mine Warfare (MIW), Mine Counter Measures (MCM), and Naval Special Warfare (NSW), particularly in the Very Shallow Water (VSW), riverine and estuarine environments. Combining the targeting and identification of different imaging data types will provides target recognition and environmental characterization with high confidence and lower false alarm rates than may be found from single source imagery.*

*The Florida Environmental Research Institute (FERI), in collaboration with the California State University System, under the NOAA-funded California Center for Integrative Coastal Observation, Research and Education (CICORE) Program, has collected nearly 15,000 square kilometers of hyper and multi-spectral imagery over a wide variety of coastal environments along the California coast. The survey sites include Humboldt Bay, San Francisco Bay, Monterey Bay, the Big Sur coast, San Luis Bay, Santa Barbara, Newport, and San Diego Harbor. These programs have generated nearly 10 TB of geo- and ortho-rectified hyperspectral image products. These image products are forming the basis for development of data fusion techniques, which will create the new imagery products that are required by military planners and civilian resource managers.*

*The ability to create new image products via data fusion techniques, as well as create education programs to train new remote-sensing scientists, requires the ability to rapidly archive and distribute high-bandwidth imagery data streams across campus-wide intranet, as well as community-wide internet services. Cal Poly has the unique campus-wide infrastructure of GigE fiber linkages between departments and computation facilities, as well as within being connected to the Internet2, with a bandwidth of >1 Gb/sec. This computation infrastructure could provide a critical backbone to the creation of new geospatial information products from the fusion of remote sensing data.*

*In addition to imaging sensor and algorithm development, FERI has created an interactive web-based data distribution site for remotely sensed data called Hyperspectral Data Repository Online (HyDRO; Figure 1; <http://hydro.marine.calpoly.edu/>; <http://www.ferihydro.org/>). The site utilizes dynamic web pages to handle user product requests, privileges, imagery access, and image sub-selection (the “front-end”). Requests made through this interface are inserted into a SQL database, which also manages job history files, user information, and priority queuing. This database is queried by an IDL-based “back-end” component that satisfies the user’s requests by parsing through the necessary imagery data files for the band selection and/or product generation for the requested regions of interest. This software system provides the rapid archival and dissemination of high bandwidth imagery files, such that the education of new scientists and the generation of new imagery products may be accomplished in fractions of the time previously spent attempting to manage and*

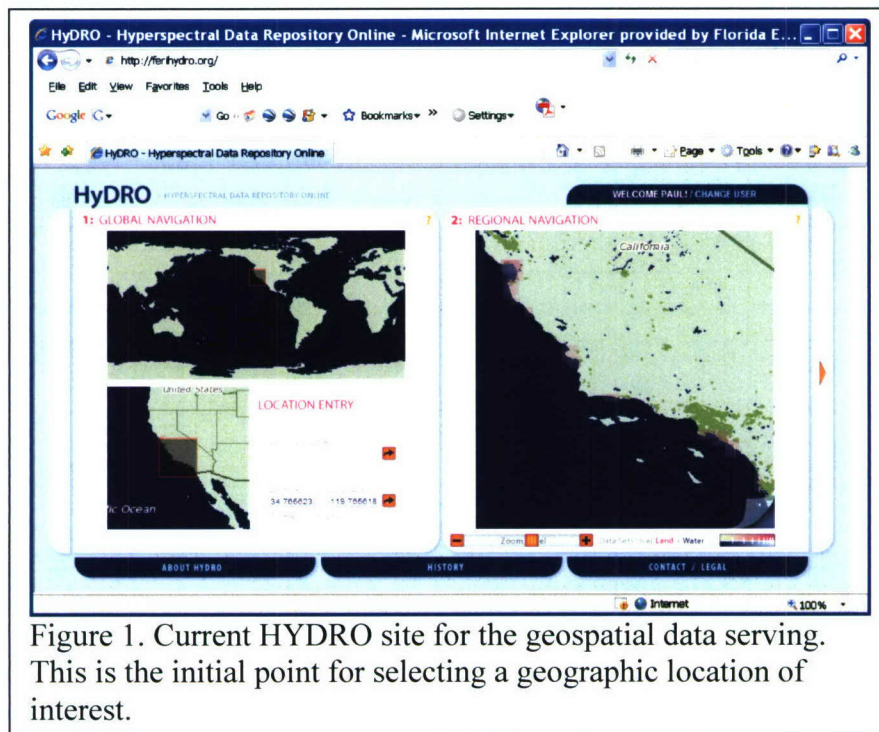


Figure 1. Current HYDRO site for the geospatial data serving. This is the initial point for selecting a geographic location of interest.

*manipulate these data.*

*This program has begun to develop a system with software, distribution, and education capabilities of high bandwidth image processing at Cal Poly, through a partnership*

between Mark Moline (Cal Poly) and Paul Bissett (FERI). This partnership provides for new capabilities being created in Cal Poly's current computational infrastructure, including the training of Cal Poly personnel on the issues of high bandwidth data management, data distribution, and image manipulation. We will build upon the current successes of the CICORE Program to provide a CSU-wide solution to the distribution and fusion of remote sensing data, and to establish Cal Poly as a hub in geospatial technology development on the West Coast.

The newest version of HyDRO (version 1.5) contains the entire original programming and functionality of the original HyDRO v 1.0, but is upgraded with several major features to provide an enhanced user experience and functionality. Major enhancements include:

- 1) v 1.5 is accessible to other web browsers besides Internet Explorer, e.g. Firefox.
- 2) Development within Linux/Apache/MySQL/PHP (LAMP) web server environment.
- 3) Single window browser operation.
- 4) Metadata display with geospatial data selection and ordering.
- 5) Enhanced search and sort capabilities for indexing by data type, data producer, location, etc.
- 6) Integration with Google Earth and NASA World Wind geo-browsers.

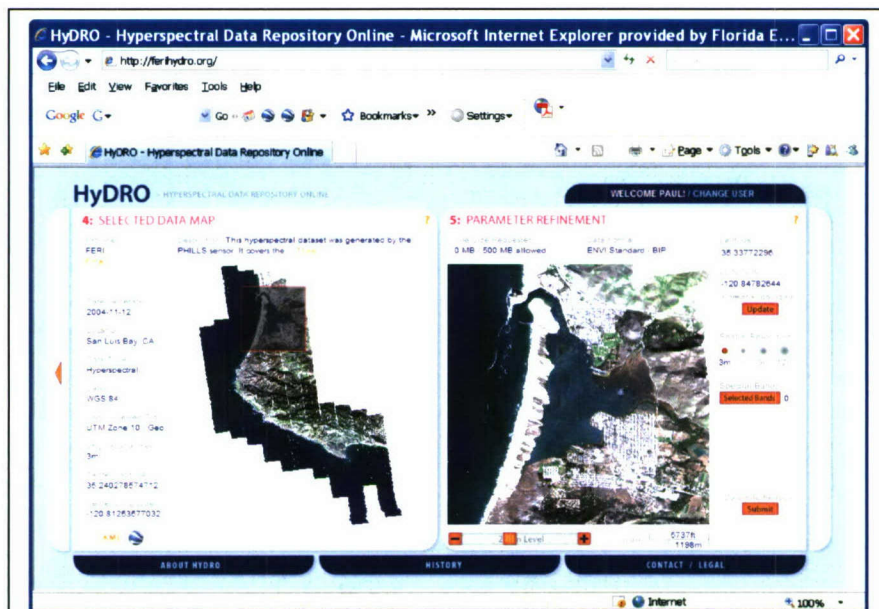


Figure 2. Refined selection from Figure 2 of the local Morro Bay. Hyperspectral data was collected by FERI as part of NOAA's CICORE program. The selection area, raster image are shown with the data preferences on the right hand side. After the user completes the specifications, the request is submitted and the data and metadata are delivered.

HyDRO v 1.0 was tailored for Microsoft's Internet Explorer for several reasons including ease of development in the .NET environment and the ubiquitous nature of

*Internet Explore (IE). However, while the greatest fraction of people accessing the HyDRO server still use IE, a growing fraction of users were requesting other browser support. In addition, the licensing cost for usage of the HyDRO application, which includes Microsoft Window Server and Microsoft SQL Server, as well as its development in the .NET environment, can be onerous for small laboratories or individual university departments. Thus, we decided to implement an open source solution (LAMP) in HyDRO v 1.5 to minimize these expenses for Cal Poly, as well as future users. We have also designed a single desktop solution that will allow the user to stay within a single browser window while performing all of the required search/sort/selection criteria. Lastly, v 1.5 will include the capability to distribute any raster mapping file, not just hyperspectral data. This will allow Cal Poly to serve other imagery types, such as satellite imagery, as well as maps produced by other means in terrestrial or aquatic environments, such as land use change maps, species and density distribution maps, bathymetric maps, and water quality maps. The next phase of the project is to incorporate in situ data collected by autonomous underwater vehicles into the system as a means of increasing this geospatial data distribution.*

#### *Technical summary*

*Initial work on the project began in February, 2006 and started with the customization of the dynamic web pages to handle the user product requirements, privileges, imagery access and image sub-selection (Figures 1 and 2). The hardware and software requirements for the Cal Poly implementation of the geospatial imagery product server were also identified and implemented. The HyDRO v 1.5 image server was installed in the Cal Poly Computational Center on September 27, 2006. The system components include:*

#### **Hardware**

*Dual core processor server (Currently being upgraded)*

*Minimum 4 GB RAM*

*Minimum 1.8 TB storage (expandability is a preferred option) in a RAID 5 configuration*

#### **Software**

*Red Hat Enterprise Linux pre-installed*

*IDL/ENVI for Linux pre-installed*

#### **IT Infrastructure**

*This server is placed in the main Cal Poly ITS Ops facility. This facility has redundant power and cooling, as well as near-line backup services. In addition, this facility is connected to the on-campus Gigabit Ethernet backbone.*

### **3. Research capacity.**

Support was also provided to build interdisciplinary capacity in the area of Radio Frequency Identification. A summary of results follows.

## ***Multidisciplinary RFID Research***

### *Investigator(s) and Department(s)*

*Tali Freed – Industrial & Manufacturing Engineering*

*Jay Singh – Industrial Technology*

*Dean Arakaki – Electrical Engineering*

*Rollin Strohman – BioResource and Agricultural Engineering*

*Jon Beckett – Animal Science*

*Michael Hall – Animal Science*

### *Abstract*

*We have found that Cal Poly's reputation, location, and alumni network generate high demand for RFID (Radio Frequency Identification), or AIT (Automatic Identification Technologies) expertise. Many of these industrial requests for assistance require multidisciplinary expertise. PolyGAIT – The Cal Poly Laboratory for Global Automatic Identification Technologies – was established as a multidisciplinary applied research and development lab created to meet the needs of the emerging RFID and other identification and tracking technologies. Thus PolyGAIT is well positioned to offer the unique set of skills that most industrial organizations do not own and cannot easily facilitate.*

*PolyGAIT faculty and students have been performing innovative, advanced RFID research and development. In order to enable these efforts financially, we sought funding collectively, as well as individually. In the 05-06 academic years we received funding from ONR via C3RP. This document consolidates results from the four RFID projects that were funded. Our goals were:*

- Position Poly GAIT as a leader of development and implementation of RFID, and as a nationwide center of academic excellence;*
- Develop Poly GAIT into a self-sustained center, generously supported by industry and government funding.*

*It is evident from this report that PolyGAIT achieved the first goal and became one of the largest and most advanced RFID labs in the country. We also achieved the self-sustainability goal by performing sponsored projects for dozens of companies and organizations and by receiving cash and equipment donations in excess of \$250,000.*

*Poly GAIT is currently in the process of becoming a university center. The Center status is expected to be granted by June 2007.*

### *Development of PolyGAIT, Tali Freed*

#### *Department of Industrial and Manufacturing Engineering, College of Engineering*

*Begun as a concept three years ago, PolyGAIT has become one of the leading RFID laboratories worldwide. (Please see [www.polygait.calpoly.edu](http://www.polygait.calpoly.edu) for more information). The PolyGAIT faculty have performed approximately 100 RFID*

projects; submitted a patent application; generated five different RFID courses (Information Systems, Supply Chain, Electronics Manufacturing, Systems Design) at the undergraduate and graduate level; involved faculty and students from five different Cal Poly colleges; developed training modules for government and industry; evaluated academic and commercial research efforts; presented at conferences; hosted representatives from academic and commercial organizations; acquired 20 different brands of readers, three types of tag printers, spectrum analyzers, GPS units, and dozens of types of tags; built a conveyor system used for testing reading reliability under various speeds and tag conditions, as well as development of RFID-based routing; and acquired computers and software, including a customized server, for software development. A partial list of the projects performed at PolyGAIT includes:

- Tagging Reusable Plastic Containers for the Fresh Produce IndustryOptimal Configuration Definition: Tags, Readers, Software, and Process for the Pharmaceutical IndustryOptimal Configuration Definition: Tags, Readers, Software, and Process for the Chemical Industry
- Asset Tracking for a Defense Electronics Company
- Computer Tracking for a Communications Company
- ROI of Automatic Data Collection in Cold Storage
- Heart-Rate and Body Temperature Monitoring
- RFID-Enabled Door Access System
- RFID-Controlled Home/Office Environment
- Event Registration and Reporting System
- RFID in the Produce Industry - Educational DVD
- Passive tag-based Local Positioning Systems for Warehouses

Several of these projects are in-line with the efforts made by the US Navy and Department of Defense to identify and track assets, as well as parts and materials along their supply chains. We are working with several major DoD and Navy suppliers (Raytheon, Lockheed-Martin, Navis, L-3 Communications, and Northrop-Grumman) to streamline their supply chains using RFID. Reports on four individual seed projects follow:

Lab Enhancement for Produce and Meat Safety Testing, Jay Singh  
Department of Industrial Technology, Orfalea College of Business

Funding was used towards the purchase of a compression test system. The funding provided by the C<sup>3</sup>RP grant was instrumental in obtaining the remainder of support from OCOB. A compression test unit was one of the major pieces of equipment needed to replicate the real world distribution environment in the validation process of consumer goods RFID applications. Following is the description of the unit:

**Lansmont Model 152-30 Compression Test System with Bench Top Controls**

The system is used to test individual packages, pallet, and unit loads. Loading platens are 60 in. by 60 in. square. Maximum machine opening is 84 in., and maximum force is

30,000 lbs. The Model 152-30 includes Lansmont's Bench Top computer-based controls and instrumentation, which are used for setting up, initiating and documenting all tests. The Model 152-30 is capable of constant load tests, constant deflection rate tests, user-defined profiles, and ramp-to-load and release tests. Data output is via color printer (included) or file transfer.

In addition a spectrum analyzer was purchased.

Several testing engagements are underway. With Navis and SAVR Communications a GPS/RFID system for container tracking is being tested. A proposal has been submitted to USDA for testing and solution development for meat tracking.

Produce Traceability Back to Harvest Field, Rollin Strohman  
Department of BioResource and Agricultural Engineering, College of Food, Agriculture and Environmental Science

In this project we are designing and building a prototype model of a process/system in which harvested produce is traceable within minutes back to the harvest field and all the transaction locations along its supply chain. The goal is to improve food safety, and prevent intentional attempts to endanger the food supply. Another goal is to increase efficiency in the produce supply chains, such that the realized cost savings well exceed the cost of implementing the process/system.

The system programs an RFID tag and records the global positioning system (GPS) location of each container on the harvester. This data is sent to a relational database where it can be associated with the harvest crew time-card data. Thus any RPC (Reusable Plastic Container) can be associated with field position and harvest crew. A National Instruments Compact FieldPoint provides the data acquisition, synchronization, and control on the harvester. The data is downloaded to a server nightly, or wirelessly transmitted to the server for online availability. Trucks moving produce from field to cooler and from cooler to customer are tracked with Trimble's Trimtrac [http://www.trimble.com/trimtrac\\_about.shtml](http://www.trimble.com/trimtrac_about.shtml). The truck locations are shown graphically with ArcGIS. Produce containers are tracked in a warehouse management system while at the facility. ArcGIS allows the database information to be shown in its geographic location. Sub-inch GPS accuracy is achieved using a Trimble GPS Base Station located on the Cal Poly campus.

Deploying RFID for Livestock Identification, Jonathon Beckett and Michael Hall  
Animal Science Department, College of Food, Agriculture and Environmental Science

As a result of the cow identified with bovine spongiform encephalopathy (BSE) on December 23, 2003, the United States Department of Agriculture (USDA) has placed a high priority on developing an animal tracking and traceability program in the US. The goal is to be able to track every animal in the US to source of origin, and identify all animals that potentially came into contact with that animal, within 48 hours. Currently, very few domestic livestock animals are individually identified. Further, the animals that

*are identified are not uniquely identified – duplicate numbers regularly occur. Further, there is no process to record these animals in a database that can be queried for animal tracking purposes.*

*While the BSE situation provides an excellent opportunity to develop and test the animal traceability protocols on a national level, the real fear is intentional introduction of foreign animal diseases as a means of bioterrorism. While prevention of such a senseless act is virtually impossible, an effective strategy to minimize deleterious spread of the disease most certainly includes animal traceability.*

*Barriers to an effective traceability program include: unique identification of each animal, cost effectiveness of the identification, ease of use and confidentiality of the records. The US National Animal Identification Program, while in principle technology neutral, is almost certainly going to be based on Radio Frequency Identification (RFID) ear tags. These tags have been available to producers for a number of years, but have enjoyed only sporadic acceptance. The traceability initiative has increased interest in this technology as a mechanism to provide a traceability infrastructure in which each animal can be traced back to farm or ranch of origin within 48 hours.*

*While the ear-tags have been available for over ten years, improvements in technology during the last 3-4 years (both half-duplex and full-duplex tags are now available) has improved the readability and retention rates of the tags. Recent studies at Cal Poly have found that one-year retention rates exceed 98.5%, and readability of tags is in excess of 99%. Therefore, one critical component of the RFID system – tags – has been to a large extent tested and deployed.*

*The other components of the system, particularly readers and software, have not been tested as thoroughly. Through this grant, we have purchased a ruggedized laptop computer and two panel RFID readers for demonstration. We also purchased several ear tags to place in animals for testing.*

*The readers, both wand and panel, were used for demonstration at the 2006 Cal Poly Beef Field Day. The objective was to demonstrate to producers the available systems, and potential costs associated with each. Further, readability of eartags in a real-time, rapid flow situation was tested in both cattle and sheep. Unfortunately, both the Farnum (self-tuning) and Allflex (manual tuning) panel readers were only able to read approximately 70% of the animals. Therefore, effort in the RFID industries must be exerted toward developing more reliable and consistent RFID readers.*

*A student was hired to develop a user-friendly software program to collect animal identification records. This program is being tested at Cal Poly's Swanton Pacific Ranch, and will soon be ready for use by commercial producers.*

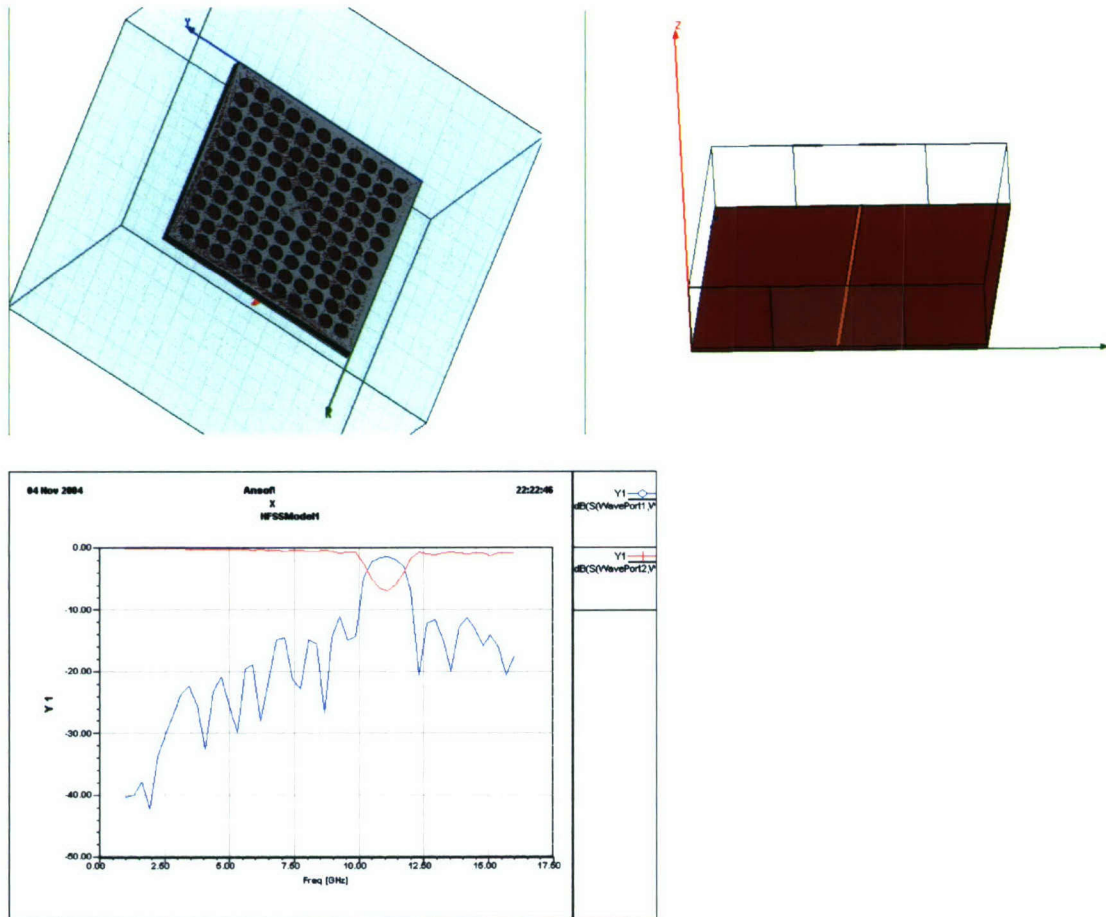
*Antenna Bandwidth Improvement and Emissions and Mounting Structure Studies, Dean Arakaki, Department of Electrical Engineering, College of Engineering*

The objectives of the project were:

- 1) Computer simulation of electronic bandgap (EBG) structures to approximate a high-impedance ground plane for mounting RFID tags to conductive surfaces.
- 2) Computer model development of fractal antenna geometries to verify multiband operation and to explore broadband design techniques to enhance performance.
- 3) Computer model to predict radiated emissions measurement accuracy inside an electromagnetic compatibility (EMC) chamber.
- 4) Computer model development of RFID mounted on flexible structures.

Project Status:

**Electronic Bandgap (EBG) Structure:** The Coccioli EBG structure [1] with 4mm diameter holes, 2.54mm substrate height, and 5.08mm periodic spacing on a 10.2 dielectric constant substrate was characterized for S-parameter performance. The EBG structure, microstrip transmission line used to verify performance, and simulated S-parameter results are shown in Fig. 1 below.

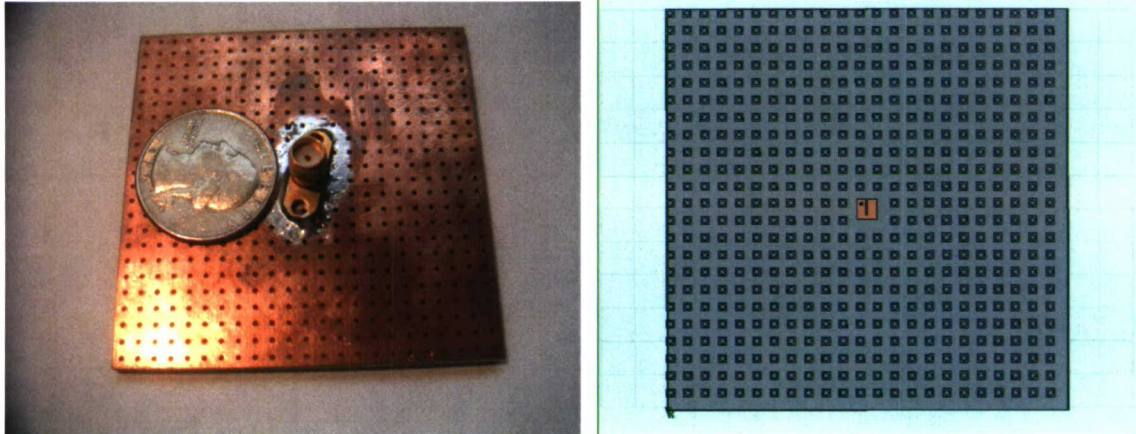


**Fig. 1:** Coccioli EBG Ground Plane Structure, Ground Plane, Microstrip Line, and S11/S21 Simulation Results

The red and blue traces in the S-parameter plot of Fig. 1 indicate  $|S_{21}|$  and  $|S_{22}|$  responses, respectively, of the microstrip line mounted on the EBG ground plane. This response indicates a 7dB attenuation of forward transmission and a reflection coefficient magnitude of -2dB at the design frequency of 11GHz. At all frequencies outside the

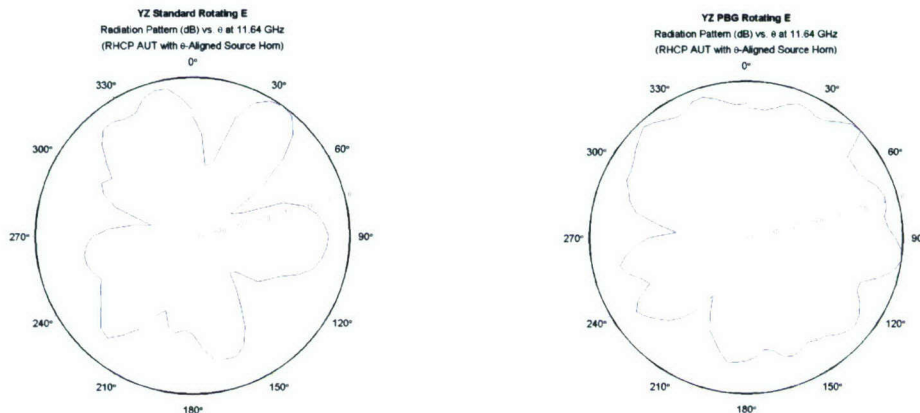
forbidden band (11GHz), the microstrip line is matched ( $|S_{11}|$  and  $|S_{22}| < -10\text{dB}$ : low reflected signal level) and transfers nearly all incident power ( $|S_{21}|$  and  $|S_{12}| > -1\text{dB}$ : low loss from input to output) to the transmit port. This confirms the desired EBG structure S-parameter characteristics for RFID ground planes.

The EBG high-impedance ground plane with an LCP (left-hand circularly polarized) slotted patch antenna was also fabricated then tested in the Cal Poly Anechoic Chamber: see Fig. 2 below.



**Fig. 2:** Electronic BandGap Ground Plane  
SMA Probe Feed and Mount for Circularly Polarized (LCP) Patch Antenna

Radiation patterns of the LCP patch antenna for the vertical polarization were measured for both the solid and EBG ground planes. The patterns indicate a suppression of surface waves as illustrated in the reduction of pattern ripple: Fig. 3.

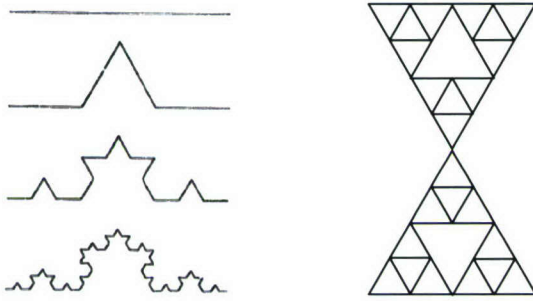


**Fig. 3:** Radiation Patterns for LCP Patch Antennas  
Mounted on Solid vs. EBG Ground Planes

With the performance of the EBG high-impedance ground plane verified for suppressed transmission along a microstrip line and surface waves at the design frequency, the EBG structure will now be applied to RFID tags. The effects of solid versus high-impedance ground planes on the performance of RFID tags will be examined using computer

modeling. The EBG structures will be refined to minimize substrate height, then fabricated on conductive surfaces to confirm simulated performance.

*Fractal Antenna Geometries: The Koch Curve Dipole, Sierpinski gasket, and Cantor array fractal patterns have been modeled and simulated using the Ansoft Designer code. The Koch curve dipole and Sierpinski gasket geometric patterns, both illustrated to three generations, are shown in Fig. 4 below.*



**Fig. 4:** Koch Curve Dipole and Sierpinski Gasket Fractal Patterns

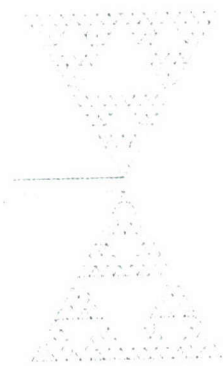
*The linear Cantor array pattern is defined by the following pattern development (again to three generations):*

*Stage 1: 101*

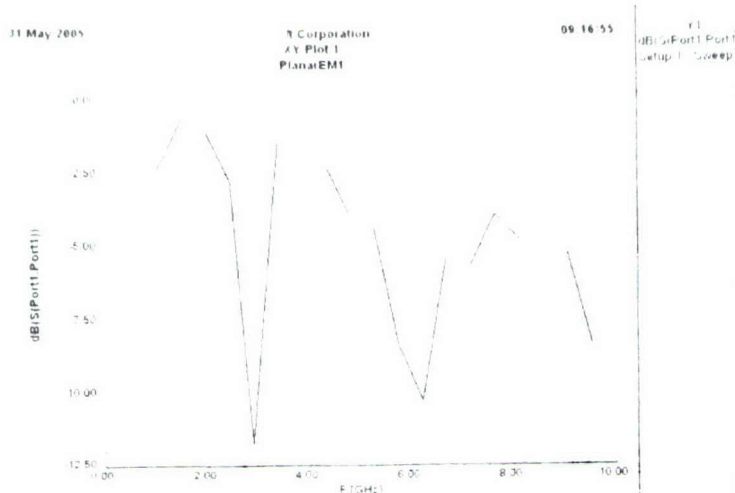
*Stage 2: 101000101*

*Stage 3: 101000101000000000101000101*

*This sequence is similar to the Koch dipole fractal pattern, except that it is along one dimension. A four-iteration Sierpinski bow-tie microstrip antenna was modeled in the Ansoft Designer code. The antenna geometry is shown in Fig. 5 below, while the input matching characteristic  $|S_{11}|$  (dB) is illustrated in Fig. 6.*



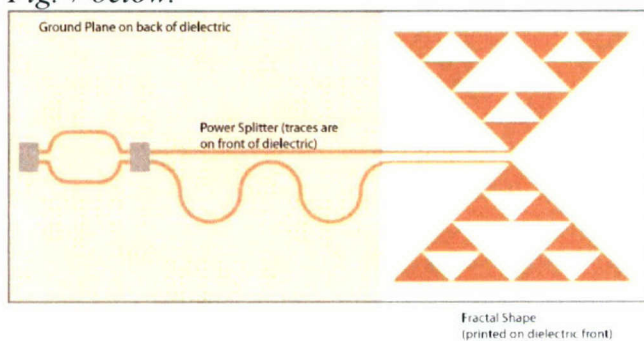
**Fig. 5:** Four-Iteration Sierpinski Fractal Bow-Tie Microstrip Antenna Geometry



**Fig. 6:** Input Matching Characteristic,  $|S_{11}|$  dB vs. Frequency (GHz)  
Four-Iteration Sierpinski Bow-Tie Microstrip Antenna

Since the scaling factor between consecutive iterations of the Sierpinski patterns of Figs. 4 and 5 is two, matched conditions are also expected at a frequency log-scale factor of two. This effect is illustrated in Fig. 6.

To obtain matched conditions at the design frequency and subsequent log-scaled frequencies, a matching network operating at the desired frequencies must be used to drive the antenna. A microstrip realization of such a feed network [2] is illustrated in Fig. 7 below.



**Fig. 7:** Microstrip Co-Planar Feed Network for Driving Fractal Geometries [2]

This structure is composed of a Wilkinson power divider (division of power into two transmission lines), followed by two microstrip lines with an electrical line length differential of  $(2n+1)\lambda/2$  ( $n = 0, 1, 2, \dots$ ). The second microstrip line provides the required 180° phase shift between the two feedlines to the fractal geometry.



*One student did preliminary work on this project, but did not make significant progress. The primary difficulty is to acquire a sufficient number of ferrite tiles (5,976) to construct the radiated emissions chamber. We obtained a quote of \$3.50 per tile from ETS-Lindgren (no university discount available), which corresponds to a total cost of \$20,916. This amount was not included in the C3RP proposal budget. We will focus on other sections of this project and pursue this portion at a later time.*

*RFID Flexible Ground Plane Structures: Will develop computer models of flexible ground planes and also incorporate the EBG high-impedance ground planes into the analysis.*

*RFID (radio frequency identification) tags cannot be mounted directly on conductive surfaces due to decreased radiation effectiveness. High-impedance ground planes can be used, but are relatively high profile and expensive to construct. As an alternative, the PIFA (planar inverted-F antenna) geometry will be explored to address both issues. Computer models will be developed and optimized, followed by fabrication of prototypes to verify computer predictions. This portion of the project will require approximately \$5,000.*

*To further decrease the size of hand-held and portable wireless devices such as cell phones and laptop computers, negative-index (left-handed) materials – also known as metamaterials – are being examined to lengthen the wavelength of substrate and subsequently reduce the required size of the microstrip patch antenna. A metamaterial substrate will be developed in a computer model, followed by the design of reduced-size microstrip patch antennas. This project will require approximately \$5,000.*

#### *References*

- 1) Coccioli, R., Itoh, T., "Design of Photonic Band-Gap Substrates for Surface Wave Suppression," *Microwave Symposium Digest, 1998 IEEE MTT-S International Symposium, Baltimore, MD., Vol. 3, Pages 1259-1262, 7-12 June 1998.*
- 2) Gianvittorio, J., "Fractal Antennas: Design, Characterization, and Applications," *M.S. thesis, University of California, Los Angeles, 2000.*
- 3) Huang, J-F., Kuo, C.W., "CPW-Fed Bow-Tie Slot Antenna," *Microwave and Optical Technology Letters, Vol. 19, No. 5, December 5, 1998.*
- 4) Ukkonen, L., Syndanheirno, L., Kivikoski, M., "A Novel Tag Design Using Inverted-F Antenna for Radio Frequency Identification of Metallic Objects," *2004 IEEE/Sarnoff Symposium on Advances in Wired and Wireless Communication, pages 91-94, 26-27 April 2004.*
- 5) Rajab, K.Z., Mittra, R., Lanagan, M.T., "Size Reduction of Microstrip Antennas Using Metamaterials," *2005 IEEE Antennas and Propagation Society International Symposium, Vol. 2B, pages 296-299, 3-8 July 2005.*

#### **4. Seed funding for new areas of research emphasis.**

Brief summaries follow of the two projects that received support.

A pilot project was funded in the area of human nutrition. The purpose of the study was to determine ways to enhance the ability of people to utilize cellulose efficiently. This would provide a more readily available energy source and a better utilization of our natural resources. The project was a collaborative effort between Dr. Rafael Jimenez of Cal Poly's Dairy Products Technology Center and Dr. John Howard of Applied Biotechnology, Inc., a small R&D company and the first company to be located on the Cal Poly campus. The project is still underway. The specific results have been omitted since they represent potentially commercializable intellectual property. The project is an initial effort for us in the area of biotechnology applied to human nutrition. We hope to expand work in this area with emphasis on human performance, including cognitive performance.

A second pilot project is described below.

***Applications of Shear-Induced, 1-D Aligned, Nano-Particles, PI: Ray Fernando***

*This summary report covers work performed during the period of October 1, 2006 – March 15, 2007. During earlier research, conducted prior to the reporting period, to understand mechanisms of mechanical property improvements offered by inorganic nanoparticles in organic coatings, it was discovered that the nanoparticles can be aligned in one-dimension (1-D) under shear deformational applications. Such well defined and stable 1-D arrangements of particles prepared under simple shear deformations have not been reported before. There is considerable interest in this field because of their potential in nanowires and other electronic, optical, and micromechanical applications. More details of this phenomenon and background of the field are available in the attached paper presented at the ACS Fall meeting in 2006.*

*There are several objectives of the specific project that this report focuses on:*

- 1. Conduct experiments with predispersed nano-silica and nano-alumina to understand the fundamental mechanisms of 1-D alignment under shear*
- 2. Learn how to make stable dispersions of electrically and magnetically conductive nanoparticles in liquid coating media*
- 3. Attempt to prepare 1-D nanowires using dispersion prepared in 2*
- 4. Attempt to measure conductivity of samples prepared in 3*
- 5. AFM and SEM characterization of nanoparticle composites*
- 6. Prepare examples for patent applications, as necessary*

*During the reporting period, a significant amount of effort was spent on further characterizing the features of 1-D aligned silica and alumina nanoparticles in clear polyurethane coatings. SEM analysis confirmed the presence of 1-D alignment of particles shown previously by AFM imaging. Supplies and parts for an optical microscope, purchased from this funding, enabled surveying of coating samples to determine areas of potential 1-D alignments that could be imaged with AFM and SEM.*

*Several additional nanomaterials were purchased. These include nano iron oxide and indium tin oxide. Efforts to achieve nano-scale dispersion of iron oxide are underway. Once dispersion is achieved, experimental work will focus on determining process and composition windows for preparing 1-D aligned iron oxide particles.*

The published article containing background information follows.

# Shear-Induced 1-D Alignment of Alumina Nanoparticles in Coatings

Lucas J. Brickweg, Bryce R. Floryancic, and Raymond H. Fernando

Polymers and Coatings Program, Department of Chemistry and Biochemistry, California Polytechnic State University, San Luis Obispo, CA 93407; www.polymerscoatings.calpoly.edu

## INTRODUCTION

Experimental and theoretical aspects of highly ordered colloidal particle assemblies have received a great deal of attention in recent years (1-9). Applications such as electronics, optics, and medical science could benefit from the unique properties of such highly ordered materials. Many of the previous studies have addressed shear-induced alignment of concentrated hard sphere suspensions. Particle size uniformity is an important factor in forming ordered structures under shear flows. Shear-induced particle string structure formation has also been reported in such dispersions, and such behavior has been correlated with shear-thinning at high rates of shear flow (1,2). The transient string formation has been shown to occur as 3-D ordered structures of particles begin to break up leaving strings of particles aligned parallel to direction of the shear flow(1).

The vast majority of prior literature in this area deals with 2-D and 3-D ordered structures. Literature on 1-D particle arrangements is limited. The most widely reported method for producing pearl-chain formations are dielectrophoresis, and other electrostatic techniques (6,7,9). The structures created using these methods are generally complicated, and often are not single, straight wires. One common goal of arranging nanoparticles in 1-D is the production of nano-scale conductive wires for electronics applications.

This paper discusses shear-induced assembly of nanoparticles into nearly perfect straight and parallel lines produced by simple application methods. The levels of nanoparticles used are much lower than the high volume fraction studied reported in previous literature. The lines produced are single aluminum oxide nanoparticles arranged in an extended pearl-necklace fashion in two-component polyurethane coatings(10). The single particle wide linear arrangements are unbranched, and hold their linearity for as long as 5 centimeters. Our literature searches so far have not uncovered reports of 1-D straight-line arrangements of particles more than a few hundred microns in length.

## EXPERIMENTAL

The alumina and silica nanoparticle containing polyurethane coatings were prepared according to the same methods described in a related study presented at this symposium(10). Commercially available alumina nanoparticle dispersions (Table 1) were incorporated into the two-component polyurethane clear coating formulation at and below levels of 1.0 wt % (i.e., approximately 0.0025 volume fraction) of the cured film.

Table 1. Nanoparticles Used in the Study

Type	Medium	% Solids	Particle Size
Alumina C	Methoxypropyl acetate with silicone additive	32.0	~25 nm
Alumina D	Propylene Glycol Ethyl Ether Acetate	32.0	~25 nm
Silica A	Methoxypropyl acetate/methoxy propanol with silicone additive	32.0	~25 nm

Coating samples were prepared on glass microscope slides with different application methods and cured at 70°C for 30 minutes. The application methods included spray, drawdown, and gentle deposition of the liquid coating drops by a dropper. These methods represent different levels of application shear rates. Coating thicknesses were in the range of 0.25 to 1.5 mils (6-38 microns).

The samples were analyzed with a Pacific Nanotechnology Atomic Force Microscope using an Ultrasharp silicon cantilever's "C" tip. Close-contact mode was used for all AFM imaging.

## RESULTS AND DISCUSSION

AFM analysis and the visual appearance of all of the coatings examined in this study indicated that, in general, the nanoparticles are well dispersed. Few areas of nanoparticle aggregates can be found on all coatings tested. AFM images of most areas on the films are similar to what is reported in reference 10, Figure 4.

**Application by Drawdown Method.** Drawdown application method produced nanoparticle strings that are very straight and continuous for a few hundred microns. AFM height images of coatings containing Alumina C (0.67 wt.%) and Alumina D (1.0 wt.%) are shown in Figure 1. These samples were prepared at an approximate application speed of 5 cm/sec. Stable, multiple straight line formations of alumina nanoparticles were observed on both coatings. In some areas, well-defined single strings of particles were found. As the application speed increased from 1 to 10 cm/sec, the abundance and the length of particle strings increased.

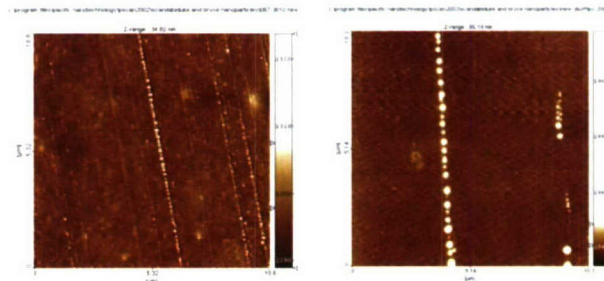


Figure 1. AFM height images (10x10 micron) of coatings containing nanoparticles: (a) Alumina-C (0.67 wt.%); (b) Alumina-D (1.0 wt.%)

**Application by Drop Deposition Method.** Four drops of the coating were placed on a glass slide, allowed to flow-out and spread for 10 minutes, and cured under the same conditions as before (70°C for 30min). No linear formations were found by the AFM analysis. Very few aggregates were visible with the optical microscope attached to the AFM, and they did not display any evidence of linear particle strings.

**Application by Spray Method.** Spray application produced the largest abundance of strings which covered the entire microscopic slide. The AFM images of this sample show many parallel strings (Figure 2) in the spray direction. Difference in sizes appears to be due to the differences in the depth of particle strings

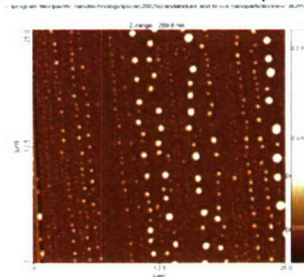
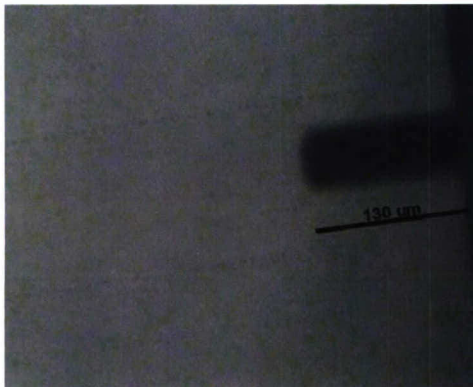


Figure 2. AFM height images (25x25 micron) of coatings containing nanoparticles: Alumina-C (0.5 wt.%)

within the film. AFM phase contrast images support this hypothesis. It is possible that the strings may be present throughout the entire thickness of the coating.

The Pacific Nanotechnology AFM used in this work utilizes a 1000x microscope with a video camera to aid in the placement of the scanning tip. Grey "streaks" can be seen on the microscope screen, and when these "streaks" or the areas surrounding them are scanned, nanoparticle strings can be found. Strings can also be seen on the microscope screen as long chains of particles that cover the entire surface of the sample. Strings are visible on the optical microscope, and have been measured to be more than 5 cm long with a 0.5 wt.% Alumina C sample prepared by spray method.



**Figure 3.** Optical micrograph of Alumina C (0.5 wt.%); prepared by spray application. Linear formations exist over the entire slide parallel to the direction of application.

**Silica Containing Coatings.** Silica A containing coatings did not show string formation with any of the application methods employed in this study.

**Effect of Shear Rate.** Estimated shear rates as low as  $13 \text{ s}^{-1}$  (0.05 cm/s at 3 mil wet coating thickness) produced strings of nanoparticles. Drop application method that involves minimal shear did not produce detectable nanoparticle strings. Efforts are underway to determine a critical shear rate for string formation. Spray method produced the largest abundance of particle strings. Although the modes of deformations in spray applications are complex (11), it is a process that involves very high shear rates.

**Effect of Nature of Nanoparticles.** As indicated above, although the two grades of Alumina particles produced the 1-D strings, the silica particles did not. It is reasonable to expect that the surface characteristics of the particles would play a critical role in this phenomenon. The two grades of alumina are reported to have highly uniform particle sizes (10). These grades resulted in coatings with better film clarity than others used in the study reported in reference 10. Alumina C, as received, contains a silicone additive, whereas Alumina D does not. No significant difference between Alumina C and Alumina D was found with respect to nanoparticle string formation in coatings.

**Effects of Particle Strings on Coating Performance.** The type of nanoparticle strings reported in this study could produce interesting visual effects in coatings. In addition, their distribution in this manner could produce interesting mechanical properties.

## SUMMARY

Atomic Force Microscopy was used to study shear-induced alignment of alumina nanoparticles in two-component polyurethane clear coatings. 1-D strings of nanoparticles, formed in an extended pearl-necklace fashion were observed near the surfaces of cured films at nanoparticle volume fractions less than 0.05. This alignment is affected by the shear conditions of the application method. When applied by spraying, linear nanoparticle strings as long as 5 centimeters were observed in the direction of shear. Nanoparticle strings were also found, to a lesser extent, when coatings were

applied by a drawdown method. The phenomenon was not observed in coatings applied with minimal shear. These particle string formations, in addition to affecting the performance of coatings, may have broader implications in the field of nanomaterials. Our literature searches so far have not uncovered reports of stable, 1-D nanoparticle arrangements with same degree of linearity produced under simple shear with compositions having very low particle loadings.

## FUTURE WORK

Further characterizations of the coatings reported in this study are underway in collaboration with the National Institute for Standards and Technology (NIST). In addition, the effects of coating compositions, viscosities, curing conditions and a wide range of other variables are also underway. Detailed particle size analysis of the alumina and silica samples is also underway.

## ACKNOWLEDGEMENTS

The authors would like to thank BYK-Chemie Company for providing the nanoparticle samples used in this study.

## REFERENCES

- 1) Ackerson, B., *J. Rheol.* **1990**, 34(4) 553-590.
- 2) Ackerson, B. *Phys. Rev. Lett.* **1981**, 46(2), 123
- 3) Cohan, I., Mason, T. G., and Weitz, D. A., *Phys. Rev. Lett.* **2004** 93(4), 04600-1.
- 4) Haw, M. D. Poon, W.C. K., and Pusey, P. N. *Phys. Rev. E.* **1998**, 57(6) 6859-6864.
- 5) Pham, H. H., Gourevich, I., Oh, J. K., Jonkman, J. E. A., and Kumacheva, E., *Adv. Mater.* **2004**, 16(9), 516
- 6) Kretschmer, R., Fritzsche, W., *Langmuir.* **2004**, 20 11797-11801
- 7) Winkleman, A., Gates, B., McCarty, L. S., and Whitesides, G. M., *Adv. Mater.* **2005**, 17, 1507-1511
- 8) Shenhar, R., Norsten, T. B., and Rotello, V. M., *Adv. Mater.* **2005**, 17(6), 657
- 9) Bhatt, K. H., and Velez, O. D., *Langmuir*, **2004**, 20, 467
- 10) Floryancic, B. R., Brickweg, L. J., Comer, J. B., Sung, L., Forster, A. M., and Fernando, R. H., *PMSE Proceedings*, **2006**, ACS National Meeting, San Francisco, CA
- 11) Xing, L.-L., Glass, J. E., and Fernando, R. H., *J. Coat. Tech.*, **1999**, 71(890), 37

### **C. Detailed research reports**

The remainder of this report contains detailed individual reports of the technical results of the research projects carried out during this project period. They are presented in the following order:

- 1. Results of CADRC research**
- 2. Results of research projects**

## IV.C.1: Results of CADRC Research

**Knowledge management using semantic web languages and technologies:  
Collaborative Agent Design Research Center (CADRC)**

Project Investigators:

Steve Gollery and Jens Pohl  
Collaborative Agent Design Research Center  
California Polytechnic State University  
San Luis Obispo, California

## Executive Overview

Every organization of more than a few persons generates large amounts of information critical to the functioning of that organization. In many cases, this mass of information is stored in separate documents, in databases accessible only through specific applications, or in the minds of the people responsible for it. All of these types of storage increase the difficulty of information retrieval and especially of constructing and discovering relationships among pieces of information stored in different media.

Especially problematic is information embedded in documents. While the contents of such documents are readily understandable by most human readers, it is difficult for software systems to extract the information and to associate that information from knowledge found in other sources. Because documents are constructed for the purpose of being read by human beings, authors are able to take advantage of the human capability to build context from natural language. This capability, unfortunately, is not available to computer software, and as a result the contents of the documents cannot be programmatically processed to generate actionable information.

To explore possible solution spaces for this problem, the Collaborative Agent Design Research Center (CADRC), with the support of the C<sup>3</sup>RP, has designed a pair of ontologies to represent two different aspects of the situation. (The word *ontology* refers to a type of information model that represents context explicitly through the relationships among the concepts in the model. Because these relationships are concrete, rather than implicit as in a human-readable document, knowledge management software can provide a level of reasoning over the information.)

The first ontology describes the organization itself at various levels of abstraction. The purpose of the organizational ontology is to enable the construction of multiple applications that have an awareness of the current state of the organization and to assist users to perform tasks within the larger organizational context. In part, the organizational ontology is intended to take the third kind of information storage mentioned in the first paragraph (the information stored only in the minds of individuals) and make that information available to software systems. This ontology also includes types of information that may be currently stored in files or databases created by standalone applications such as spreadsheets or schedulers. Because these files can only be accessed by the programs that created them, they form islands of data that cannot be programmatically linked to discover potential problems such as scheduling bottlenecks. By relating these types of information using a single ontology, the contents of these former islands become linked and accessible to organization-wide knowledge management systems.

The second ontology describes the contents of documents. The purpose of the document ontology is to allow the construction of self-describing documents. It should be possible to create documents that can be viewed by humans as if they were simple word-processing documents, yet still have a structure that is amenable to software-based reasoning. One key element of the document ontology is the concept of a document *fragment*, which may be as small as a sentence or as large as is needed to convey a single topic. Another is a set of concepts that together convey the *structure* of a document – that is, the way that a set of document fragments may be organized to construct a complete document. Because the *contents* of a document are defined separately from the *structure* of the document, it becomes possible to reuse document fragments in multiple documents, thereby preventing rework and increasing the organization's productivity.

While we believe that the document ontology holds much potential for the future, we are also aware that its adoption in any organization depends on the quality of the tools available for building document fragments, for constructing documents from many fragments, and for locating document fragments that are related to a given set of concepts. Creating these tools will be a complex endeavor. In order to gain experience in building toolsets based on these ontologies, as well as to create immediately practical applications, the efforts of the Semantic Web Knowledge Management Project over the 2005-2006 time frame have focused on the use of the organizational ontology within a set of tools.

## **Project Description**

In previous project segments, we have developed an ontology (or “knowledge model”) that can be used to represent an organization, the projects the organization undertakes, and the people working for the organization. The organization model specifically includes features that allow us to model the CADRC as an exemplar, but we believe that the model itself is general enough to be applied to many organizations (although it is likely that other kinds of organizations would require some extensions to the current ontology).

In a separate but linked model, we describe types of information objects that an organization may produce. One of the key concepts of the information objects model is that the information most often found currently in various documents has wider potential usage than the documents themselves. Following this idea, our model separates the content of the documents from the structure of the documents. The model defines a document as a specific sequence of information fragments (relatively small, stand-alone pieces of text that may range from a sentence to a few paragraphs). These information fragments may also be annotated by being associated with one or more topics, and the topics themselves may be organized by being associated with other topics in various ways.

The ultimate project goal is to produce a set of software tools that will be useful in the following areas:

1. Creating information fragments.
2. Annotating fragments, either by hand or through analysis by software.
3. Defining document structures for recurring types of documents, such as reports that must be filed on a regular basis.
4. Creating documents on the fly by specifying a structure that can be filled in by the system with candidate fragments.

## **Progress**

Since Fall, 2005, we have been at work on building an infrastructure to support the creation of systems and tools in the above areas. The infrastructure currently includes components that bridge between the ontologies and the user interface, and are divided as follows:

1. Ontology interface library. This provides a high-level abstraction over a more general ontology representation library, allowing other components to take a more application-centric view of the model. The ontology interface also allows other

components to register their interest in changes or additions to the information known to the system that is related to specific parts of the model. This registration process simplifies later access for the purpose of producing reports.

2. Report structure processor. The structure of each current report is defined by an XML<sup>1</sup> document. The report structure processor interprets these documents, retrieves information from the repository, and organizes the presentation of the information. The output of the report structure provider is a set of value providers (see the next bullet point) associated with a set of user interface elements (bullet point 4).
3. Value providers. This is a set of classes that are used in conjunction with the user interface elements (see the next bullet point) to link information elements in the repository to the visual display of the report.
4. User interface elements. These are classes that provide generic versions of some common visual display objects that are familiar from windowed computing environments. They are used in our system to avoid writing directly to a specific user interface library. Each user interface element is associated (in a loosely-coupled way) with a specific value provider, from which the user interface element retrieves a piece of information from the information repository.
5. User interface producer. This set of classes is specific to a given concrete user interface library. The main class in this component is responsible for interpreting the user interface elements produced by the report structure processor, and for constructing equivalent concrete user interface elements for the display of the report.

These components have been carefully designed so that the system can be moved to other platforms with a minimum of changes. Currently, the system uses the Eclipse Rich Client Platform<sup>2</sup> as a framework. The Eclipse RCP uses its own user interface library, the Simple Widget Toolkit<sup>3</sup>. In the knowledge management project, all references to SWT are isolated in the user interface provider (bullet point 5, above). To move the system to the standard Java user interface, it would suffice to replace the current user interface provider with one that creates Swing components.

It should also be noted that none of these components are specific to the ontologies that we have defined. Many of them would work with any model defined as an OWL<sup>4</sup> ontology, and some (the user interface components) make no assumption about the source of the text they present to the user. Only the XML documents describing report structure contain identifiers and properties that are part of the models. This design concern is central to our approach, since we assume that the current ontology will be extended or possibly will be changed altogether during the course of

---

<sup>1</sup> Bray, T., J. Paoli, C.M. Sperberg-McQueen, E. Maler, F. Yergeau, eds. Extensible Markup Language (XML) 1.0 (Fourth Edition). <http://www.w3.org/TR/2006/REC-xml-20060816/>

<sup>2</sup> Rubel, Dan. "The Heart of Eclipse", Dan Rubel  
<http://www.acmqueue.org/modules.php?name=Content&pa=showpage&pid=425>

<sup>3</sup> Northover, Steve and Mike Wilson. SWT: The Standard Widget Toolkit. Addison Wesley 2004

<sup>4</sup> OWL Web Ontology Reference, <http://www.w3.org/TR/owl-ref/>

implementation. Keeping the ontology out of the code allows us to use the same code for multiple ontologies.

### **Current Capabilities**

So far, we have demonstrated the utility of the infrastructure described in the previous section by defining a number of input forms and report types. The present system allows the definition of new projects, the addition of new personnel to the repository, and the creation of information fragments that collectively represent status reports on the work done by each person working on a given project. The system also generates collected reports by project, by person, and by date. All forms and reports are defined by report structure documents, and new reports can be added with little or no change to existing code.

We have shown that the organizational ontology can support a rich and extensible set of software tools. We have also demonstrated an approach to tool building that separates the software into three separate, loosely connected parts. At one end are those aspects of the system that are specific to a given ontology. At the other end are the aspects that are specific to a given user interface. Between are representations of information and user interface elements that can be used with any ontology and any type of user interface. Our system design approach has shown that the bulk of the system can be neutral to both ontology and user interface, and therefore can be reused in other systems.

### **Future Work**

One important aspect of moving towards the document assembly process as described in our document ontology is that the users of a system based on this ontology must have a shared conceptualization of the body (or bodies) of knowledge on which any constructed documents will be based. We will now begin to explore techniques for enabling groups of users to build this shared conceptualization incrementally and over time.

The goal of the Semantic Web Knowledge Management Project for 2006-2007 is to design and implement a software system that allows people to construct a shared model of a domain of knowledge. This model will include information within that domain, as well as categories of knowledge and the relationships between those categories. Users will modify and extend their collective model as they enter new information. The system will be able to use the evolving model to reason about the information and to discover new relationships and correlations.

The system will be built using Semantic Web<sup>5</sup> concepts, technologies, and languages. The semantic web approach to information architecture relies on the use of modeling languages with well-defined formal semantics that serve as the basis for software tools that provide automated reasoning about data and information.

Under the guidance of experienced software engineers, graduate and undergraduate students will design, implement, and test the system as it evolves over the course of the project.

---

<sup>5</sup> Shadbolt, Nigel, Wendy Hall, Tim Berners-Lee. The Semantic Web Revisited. IEEE Intelligent Systems, May/June 2006 pp. 96-101

## **Significance**

Question to be researched or studied: Will a collaborative networked environment for sharing information using semantic web technologies allow discovery of useful and actionable relationships?

Context for the problem: Information sharing among individuals and organizations is frequently hampered by the lack of machine-processable relationships between separate pieces of information. Such relationships may exist within the mental models of the human beings who enter data into a computer system, but unless there is a representation of these relationships that can be processed by software, the system will never be able to discover correlations among individual pieces of data. In a world of ever-increasing volumes of data, only automated approaches can hope to succeed in examining all possible connections, but the absence of a deep representation of the context of data limits the approaches that can be used by software.

Many attempts to provide context for data have relied on predefined models into which all data must fit. Although many systems based on such models have provided useful functionality, the type of processing and reasoning that can be done in these systems is limited by the fact that the model does not change once the system has been deployed. This condition does not match well with real-world experience, where gathering information inevitably leads to modifications of our mental model. A software system that seeks to provide context for ongoing data input must have a flexible model of an area of knowledge that can grow and change as new information and new requirements are identified by the humans using the system.

We propose to create a set of software tools that will allow users to collaboratively create a shared model of the information space representing a domain of knowledge. Users will also enter information into the system, which will categorize and relate that information based on the model. We will show that such a system allows the gradual construction and extension of a shared view of the domain, and that the system's capabilities can expand with the model.

## **Design and Procedures**

We will begin by surveying existing applications (commercial and open source) that provide collaborative workspaces. From this survey and our own understanding of the functionality that our system must perform, we will define initial requirements.

Based on those requirements, we will design user interface concepts, underlying application models, and the behavior of tools within the system.

Additionally, an initial domain model will be created. This model will be used to seed the system and provide a foundation for later expansion and evolution as more information is added to the knowledge base. Our goal is for the system itself to have no dependency on the specific domain model, so that the system can be instantiated at any time as a knowledge base for another domain. The proposed system can be seen as a shell or framework for constructing, populating and querying models in any field of knowledge.

The major effort in this project will be in the implementation of the software and the adjustment of requirements and design as the system comes online. As with any software project based on exploration and experimentation, our understanding of what is useful and important in a collaborative environment is highly likely to change once users are able to interact with the

initial version of the system. In anticipation of this reality, we will design our process to expect iteration through the requirements gathering, design, and implementation phases.

### **Conclusion**

For the past four years, the Semantic Web Knowledge Management Project has explored the use of semantic web technologies and languages in modeling organizational knowledge and providing sets of tools for creating, manipulating, and accessing knowledge organized within those models. With our latest system, we have demonstrated the validity of this approach when combined with current best practices in system architecture. This work has laid the foundations for our next effort, which will allow a community of users to model their own domain of knowledge in an evolutionary process.

### **Publications and Presentations Related to Project**

#### **Full publications**

Copies of full publications follow.

## A Framework for Truth Maintenance in Multi-Agent Systems

Brett Bojduj<sup>1</sup>, Ben Weber<sup>1</sup>, and Dennis Taylor<sup>2</sup>

<sup>1</sup> California Polytechnic State University

San Luis Obispo, CA 93407

{bbojduj, bgweber}@calpoly.edu

<sup>2</sup> CDM Technologies, Inc.

2975 McMillan Ave. Suite 272

San Luis Obispo, CA 93401

djtaylor@cdmtech.com

**Summary.** Maintaining logical consistency in a knowledge base that receives asynchronous events from collaborative agents poses a challenge, due to multiple agent perspectives of the knowledge base. Facades and filters further complicate this problem by distorting the definition of consistent knowledge. Therefore, maintaining logical consistency of the knowledge base requires an infrastructure for handling truth maintenance. This paper presents a generic, object-oriented framework for truth maintenance in collaborative multi-agent systems. The core of the framework is an agent that autonomously reasons on system events, thus guaranteeing the integrity of the knowledge base independent of external agents. Specialization to a particular domain is achieved through the description of tests that verify the consistency of the knowledge base. This paper shows an example of this approach in a real-world, multi-agent system and discusses performance and maintainability in such a system.

Key words: truth, maintenance, framework, collaborative, agent, system

### 1 Introduction

A challenge for large software systems is maintaining logical consistency between collaborative agents operating asynchronously to solve a problem. Multi-agent systems may use a divide-and-conquer approach to solve a problem [4]. They can do this by assigning different aspects of the problem to different agents. Each agent in the system may have a different perspective of the problem and the definition of consistency may vary between agents. Therefore, multi-agent systems require that solutions are consistent across all agent perspectives. This paper presents an external truth maintenance agent as a technique for maintaining consistency between collaborative agents.

Multi-agent systems have been used to solve constraint satisfaction problems [7]. For large, highly-constrained problems, implementation of an agent that simultaneously solves all constraints becomes impractical. A common solution is to construct collaborative agents that subdivide the problem into smaller problems and share solutions. Decomposing the problem requires that solutions shared between two agents are consistent with both agents' perspectives of the problem. Consistency can be achieved if agents communicate through a shared knowledge base that is monitored by an external truth maintenance agent. The truth maintenance agent verifies that solutions in the knowledge base are consistent with respect to all constraints in the problem. This alleviates individual agents from the burden of maintaining consistency for the entire solution.

This paper explores the use of a truth maintenance agent to maintain logical consistency of a knowledge base in a global logistics and scheduling application. The task of planning is distributed between several collaborative agents and the consistency of the knowledge base is verified by an external truth maintenance agent. The presented technique enables metaheuristic-based agents to generate consistent solutions in a near real-time planning environment.

## 2 Truth Maintenance Systems

Most agents require logically consistent information in order to perform their reasoning functions. In dynamic environments, additional information is added to the problem definition while the system is running, which may conflict with current solutions. If logically inconsistent data enters a system, existing solutions must be marked as invalid or revised in order to allow agents to reason accurately.

Doyle proposed a truth maintenance system to handle cases when logically inconsistent data enters a system [2]. Each set of data in the system is annotated with a node that specifies the consistency of the data. When new information enters the system, every node is updated to reflect the new problem definition. Huhns and Bridgeland extended the truth maintenance system to enable consistency in multi-agent systems [5]. When new information enters the system, every agent verifies that its shared data is consistent. This validation process becomes a bottleneck, but is required for blackboard architectures. If the knowledge base becomes compromised, then agents in the system communicating through the blackboard will be unable to properly perform their reasoning tasks.

Traditional truth maintenance systems guarantee consistency of a knowledge base, but are unsuitable for real-time systems. Elkan proved that truth maintenance systems that deal with logically inconsistent information are NP-complete [3]. Therefore, systems must compromise global consistency in order to operate in real-time environments. The amount of computation required to

enforce consistency can be reduced if agents consider different aspects of the problem. Bertel et al. presented the concept of “aspect filters” which can be applied to truth maintenance systems [1]. When new data enters the system, consistency is maintained within only the affected aspects of the knowledge base. If an aspectualization approach is used, then agents working on different aspects of a problem are not required to be logically consistent with respect to each other. This technique lends itself well to metaheuristic-based agents that perform reasoning with a shared knowledge base.

### 3 Definition of Consistency

The purpose of a truth maintenance system is to verify the consistency of a knowledge base. We define the consistency of a knowledge base,  $C_{KB}$ , to be:

$$C_{KB} = F_V(O, C) \quad (1)$$

where  $O$  is the set of objects in the problem definition and  $C$  is the set of defined constraints. The verification function,  $F_V$ , checks that all constraints are satisfied by the set of objects. The knowledge base is defined as consistent when all constraints in the problem definition are satisfied. In our implementation, we apply this definition of consistency to a real-time vehicle routing and scheduling problem.

### 4 Truth Maintenance Agent

This paper presents a technique for maintaining consistency through the use of an agent that runs independent of the rest of the system. The purpose of the truth maintenance agent is to verify that solutions shared in a knowledge base are globally consistent. Additionally, the agent checks the consistency of the knowledge base when new data enters the system. The truth maintenance agent alleviates other agents from checking the validity of objects in the knowledge base. This design facilitates communication between agents operating in a distributed environment by enabling agents with different perspectives to communicate consistent solutions.

When a solution is posted to the knowledge base, the truth maintenance agent checks the validity of the solution. During this process, other agents in the system may utilize the solution. If the solution is determined to be consistent, then the truth maintenance agent performs no further actions. However, if the solution is logically inconsistent, then the truth maintenance agent marks the solution as invalid and broadcasts an alert to the other agents in the system. Agents operating on invalid solutions stop their reasoning processes and check for revised solutions from the knowledge base or generate new solutions.

The truth maintenance agent verifies the consistency of solutions using an aspect-oriented approach. The verification function is composed of a series of tests, which check various aspects of objects in the knowledge base. The agent contains a mapping of external events to tests that specifies which tests to run when new data enters the system. This approach allows the truth maintenance agent to maintain consistency without exhaustively testing the knowledge base. However, an exhaustive approach is used when new solutions are posted to the knowledge base, because the truth maintenance agent must verify that new solutions are globally consistent.

The truth maintenance agent provides a technique for maintaining consistency in a system, but there are several tradeoffs. The agent runs independently of the system, which may result in a delay between the time when a solution is posted and the time when the validity of the solution is verified. Also, the truth maintenance agent does not monitor the internal states of other agents in the system. Therefore, the main disadvantage is that agents in the system may temporarily operate on invalid data. However, several benefits are gained through the use of a truth maintenance agent. Agents in the system are not required to translate between different perspectives, reducing code complexity. The truth maintenance agent can also be applied to dynamic problem domains. Agents continue solving a problem when a change occurs and are informed by the truth maintenance agent if the knowledge base becomes inconsistent.

## 5 Implementation in TRANSWAY

The truth maintenance agent was developed to facilitate communication between collaborative agents in a logistics and scheduling application called TRANSWAY, developed by CDM Technologies, Inc. The purpose of TRANSWAY is to provide execution plans for military distribution of supplies at strategic and operational levels of planning. The problem is represented as a constrained vehicle routing problem with multiple types of conveyances and transshipments. The main challenge for the TRANSWAY system is providing up-to-date plans, because new data frequently enters the system.

The task of planning is distributed between three agents in TRANSWAY: the routing agent, the planning agent, and the impediment agent. The routing agent maintains the shortest paths between all locations in a scenario and enforces conveyance-range constraints. The planning agent uses a tabu search algorithm to generate solutions using paths from the routing agent and handles scheduling constraints. The impediment agent verifies that solutions discovered by the planning agent do not require conveyances to travel through impediments, such as sandstorms. A multi-agent approach enables decomposition of the problem, but prevents individual agents from determining if solutions are globally consistent.

The most common usage of the truth maintenance agent is maintaining consistency when new information about impediments enters the system. If a weather impediment is created or modified, the truth maintenance agent performs the tests mapped to these events to determine the validity of the knowledge base. There are two ways an impediment can invalidate a solution: a light impediment limits the cruising speed of vehicles and results in scheduling conflicts in the solution, while a severe impediment causes routes to become impassible.

When a weather impediment is created or modified, the truth maintenance agent checks that conveyances in the solution are not scheduled to traverse through impeded routes. If a conveyance is scheduled to pass through an impeded route that makes its scheduled delivery impossible, the truth maintenance agent marks the delivery as invalid and sends an alert to the routing and planning agents. The routing and planning agents are responsible for generating a new solution that is consistent with the updated scenario definition. The routing agent validates the solution by generating routes that traverse around the impediment. Afterwards, the planning agent removes the invalid delivery and reschedules the delivery using updated routes from the routing agent.

The truth maintenance agent has enabled the TRANSWAY system to solve logistics problems with several constraints. Computationally expensive processes required to verify the consistency of solutions are delegated to the truth maintenance agent, allowing other agents in the system to continue solving a scenario. Implementation of the truth maintenance agent in TRANSWAY has demonstrated the feasibility of maintaining consistency between collaborative agents in a near real-time planning environment.

## 6 Conclusion and Future Work

Maintaining consistency of a knowledge base in a multi-agent system is computationally expensive. This process becomes impractical when agents must interact in a real-time environment. Therefore, it is necessary to tradeoff global consistency for improved performance. We have presented a technique for maintaining consistency of a knowledge base using an external truth maintenance agent. The truth maintenance agent facilitates communication between agents with different perspectives and enables agents to operate in near real-time environments.

The main drawback of the presented truth maintenance technique is that agents may operate on invalid information for large periods of time. Therefore, agents should not be completely reliant on the truth maintenance agent for validation of the knowledge base. Taylor et al. presented a technique for generating consistent solutions to dynamic problems for a single metaheuristic-based agent [6]. Future work includes leveraging this technique with the truth

maintenance agent to achieve improved agent performance in highly dynamic environments.

## References

1. Bertel S, Freksa C, Vrachliotis G (2004) Aspectualize and conquer in architectural design. In: Gero JS, Tversky B, Knight T (eds) *Visual and Spatial Reasoning in Design III*. Key Centre of Design Computing and Cognition; University of Sydney 255–279
2. Doyle J (1979) A truth maintenance system. *Artificial Intelligence* 12(3):231–272
3. Elkan C (1990) A rational reconstruction of nonmonotonic truth maintenance systems. *Artificial Intelligence* 43(2):219–234
4. Ephrati E, Rosenschein JS (1994) Divide and conquer in multi-agent planning. *AAAI '94: Proceedings of the twelfth national conference on artificial intelligence* 1:375–380
5. Huhns MN, Bridgeland DM (1991) Multiagent truth maintenance. *IEEE Transactions on Systems, Man, and Cybernetics* 21(6):1437–1445
6. Taylor D, Weber B, Bojduj B (2006) A tabu search framework for dynamic combinatorial optimization problems. *Proceedings, Integrated Design and Process Technology, IDPT-2006* 663–668
7. Yokoo M (2001) *Distributed constraint satisfaction: foundations of cooperation in multi-agent systems*. Springer-Verlag, London

# Elements of Human Decision-Making

Jens Pohl, Ph.D.

Executive Director, Collaborative Agent Design Research Center (CADRC)  
California Polytechnic State University (Cal Poly)  
San Luis Obispo, California, USA

## Abstract

The purpose of this paper is to present some understandings of the human problem-solving activity that we have gained in the Collaborative Agent Design Research Center (CADRC) over the past two decades. Since we feel strongly that the human decision-maker should be an integral component of any computer-based decision-support system, it follows that we would have endeavored to incorporate many of the elements that appear to be important to the user in the design of these systems. The complexity of the human cognitive system is evidenced by the large body of literature that describes problem-solving behavior and the relatively fewer writings that attempt to provide comprehensive explanations of this behavior. Our contributions in this field are confined to the identification of important elements of the problem-solving activity and exploration of how these elements might influence the design of a decision-support system.

## Keywords

agents, analysis, communication, computers, context, data, data-centric, decisions, decision-making, decision-support, design, evaluation, information, information-centric, intuition, problem-solving, reasoning, representation, synthesis, visualization

## Some Human Problem Solving Characteristics

Human beings are inquisitive creatures by nature who seek explanations for all that they observe and experience in their living environment. While this quest for understanding is central to our success in adapting to a changing and at times unforgiving environment, it is also a major cause for our willingness to accept partial understandings and superficial explanations when the degree of complexity of the problem situation confounds our mental capabilities. In other words, a superficial or partial explanation is considered better than no explanation at all. As flawed as this approach may be, it has helped us to solve difficult problems in stages. By first oversimplifying a problem we are able to develop an initial solution that is later refined as a better understanding of the nature of the problem evolves. Unfortunately, now we have to contend with another characteristic of human beings, our inherent resistance to change and aversion to risk taking. Once we have found an apparently reasonable and workable explanation or solution we tend to lose interest in pursuing its intrinsic shortcomings and increasingly believe in its validity. Whether driven by complacency or lack of confidence, this state of affairs leads to many surprises. We are continuously discovering that what we believed to be true is only partly true or not true at all, because the problem is more complicated than we had previously assumed it to be.

The complexity of problems faced by human society in areas such as management, economics, marketing, engineering design, and environmental preservation, is increasing for several reasons. First, computer-driven information systems have expanded these areas from a local to an

increasingly global focus. Even small manufacturers are no longer confined to a regionally localized market for selling their products. The marketing decisions that they have to make must take into account a wide range of factors and a great deal of knowledge that is far removed from the local environment. Second, as the net-centricity of the problem system increases so do the relationships among the various factors. These relationships are difficult to deal with, because they require the decision-maker to consider many factors concurrently. Although the biological operation of the human brain is massively parallel, our conscious reasoning processes are sequential. Simply stated, we have difficulty reasoning about more than two or three variables at any one time. Third, as the scope of problems increases decision-makers suffer simultaneously from two diametrically opposed but related conditions. They tend to be overwhelmed by the sheer volume of data that they have to consider, and yet they lack information in many specific areas. To make matters worse, the information tends to change dynamically in largely unpredictable ways

It is therefore not surprising that governments, corporations, businesses, down to the individual person, are increasingly looking to computer-based decision-support systems for assistance. This has placed a great deal of pressure on software developers to rapidly produce applications that will overcome the apparent failings of the human decision-maker. While the expectations have been very high, the delivery has been much more modest. The expectations were simply unrealistic. It was assumed that advances in technology would be simultaneously accompanied by an understanding of how these advances should be applied optimally to assist human endeavors. History suggests that such an a priori assumption is not justified. There are countless examples that would suggest the contrary. For example, the invention of new materials (e.g., plastics) has inevitably been followed by a period of misuse. Whether based on a misunderstanding or lack of knowledge of its intrinsic properties, the new material was typically initially applied in a manner that emulated the material(s) it replaced. In other words, it took some time for the users of the new material to break away from the existing paradigm. A similar situation currently exists in the area of computer-based decision-support systems.

### **The Rationalistic Tradition**

To understand current trends in the evolution of progressively more sophisticated decision-support systems it is important to briefly review the foundations of problem solving methodology from an historical perspective. Epistemology is the study or theory of the origin, nature, methods and limits of knowledge. The dominant epistemology of Western Society has been technical rationalism (i.e., the systematic application of scientific principles to the definition and solution of problems).

The rationalistic approach to a problem situation is to proceed in well defined and largely sequential steps as shown in Figure 1: define the problem; establish general rules that describe the relationships that exist in the problem system; apply the rules to develop a solution; test the validity of the solution; and, repeat all steps until an acceptable solution has been found. This simple view of problem solving suggested a model of sequential decision-making that has retained a dominant position to the present day. With the advent of computers it was readily embraced by 1st Wave software (Figure 2) because of the ease with which it could be translated into decision-support systems utilizing the procedural computer languages that were available at the time.

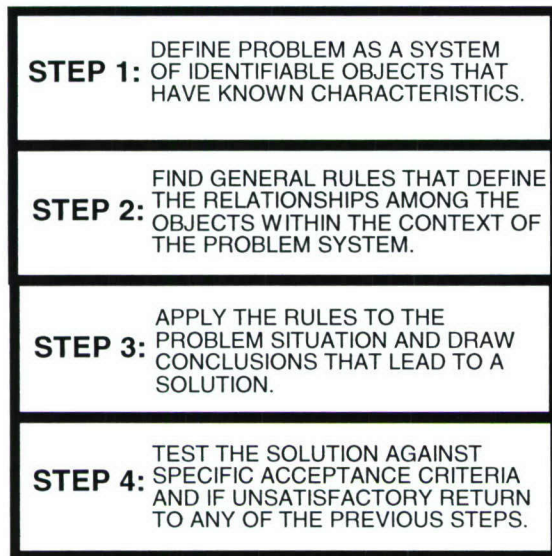


Figure 1: Solution of simple problems

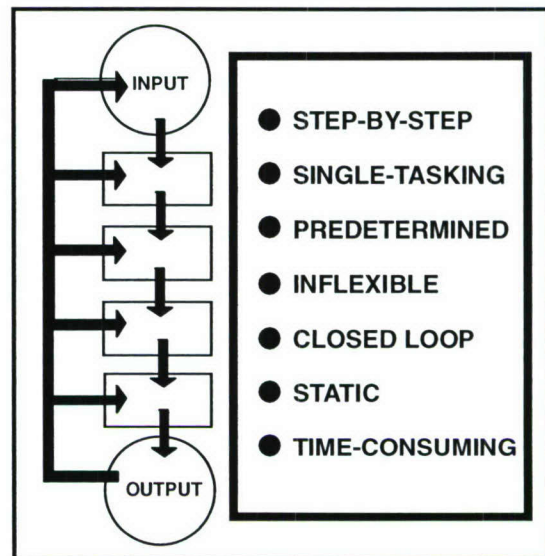


Figure 2: Sequential decision-support

The close correlation between the rationalistic approach and what is commonly referred to as the scientific method, is readily apparent in the series of basic steps that are employed in scientific investigations: observe the phenomenon that requires explanation; formulate a possible explanation; develop a method capable of predicting or generating the observed phenomenon; interpret the results produced by the method; and, repeat all steps until an acceptable explanation of the observed phenomenon has been found. Scientific research typically attempts to establish situations in which observable actions (or reactions) are governed by a small number of variables that can be systematically manipulated. Every effort is made to keep the contrived situation simple, clear and deterministic, so that the results of the simulation can be verified.

However, natural phenomena and real world problems are often very complex involving many related variables. Neither the relationships among the variables nor the variables themselves are normally sufficiently well understood to provide the basis for clear and comprehensive definitions. In other words, problem situations are often too complex to be amenable to an entirely logical and predefined solution approach. Under these circumstances the analytical strategy has been to decompose the whole into component parts, as follows:

- ◆ Decompose the problem system into sub-problems.
- ◆ Study each sub-problem in relative isolation, using the rationalistic approach (Figure 1). If the relationships within the sub-problem domain cannot be clearly defined then decompose the sub-problem further.
- ◆ Combine the solutions of the sub-problems into a solution of the whole.

Underlying this problem-solving strategy is the implicit assumption that an understanding of parts leads to an understanding of the whole. Under certain conditions this assumption may be valid. However, in many complex problem situations the parts are tightly coupled so that the behavior of the whole depends on the interactions among the parts rather than the internal characteristics of the parts themselves (Bohm 1983, Senge 1993). An analogy can be drawn with

the behavior of ants. Each ant has only primitive skills, such as the ability to interpret the scent of another ant and the instinctive drive to search for food, but little if any notion of the purpose or objectives of the ant colony as a whole. In other words, an understanding of the behavior of an individual ant does not necessarily lead to an understanding of the community behavior of the ant colony of which the ant is a part.

Decomposition is a natural extension of the scientific approach to problem solving and has become an integral and essential component of rationalistic methodologies. Nevertheless, it has serious limitations. First, the behavior of the whole usually depends more on the interactions of its parts and less on the intrinsic behavior of each part. Second, the whole is typically a part of a greater whole and to understand the former we have to also understand how it interacts with the greater whole. Third, the definition of what constitutes a part is subject to viewpoint and purpose, and not intrinsic in the nature of the whole. For example, from one perspective a coffee maker may be considered to comprise a bowl, a hotplate, and a percolator. From another perspective it consists of electrical and constructional components, and so on.

Rationalism and decomposition are certainly useful decision-making tools in complex problem situations. However, care must be taken in their application. At the outset it must be recognized that the reflective sense (Schon 1983) and intuition of the decision-maker are at least equally important tools. Second, decomposition must be practiced with restraint so that the complexity of the interactions among parts is not overshadowed by the much simpler behavior of each of the individual parts. Third, it must be understood that the definition of the parts is largely dependent on the objectives and knowledge about the problem that is currently available to the decision-maker. Even relatively minor discoveries about the greater whole, of which the given problem situation forms a part, are likely to have significant impact on the purpose and the objectives of the problem situation itself.

### **Decision Making in Complex Problem Situations**

As shown in Figure 3, there are several characteristics that distinguish a complex problem from a simple problem. First, the problem is likely to involve many related issues or variables. As discussed earlier the relationships among the variables often have more bearing on the problem situation than the variables themselves. Under such tightly coupled conditions it is often not particularly helpful, and may even be misleading, to consider issues in isolation. Second, to confound matters some of the variables may be only partially defined and some may yet to be discovered. In any case, not all of the information that is required for formulating and evaluating alternatives is available. Decisions have to be made on the basis of incomplete information.

Third, complex problem situations are pervaded with dynamic information changes. These changes are related not only to the nature of an individual issue, but also to the context of the problem situation. For example, a change in wind direction during a major brushfire may have a profound impact on the entire nature of the relief operation. Apart from precipitating an immediate re-evaluation of the firefighting strategy, it may require the relocation of firefighters and their equipment, the replanning of evacuation routes, and possibly even the relocation of distribution centers. Certainly, a change in the single factor of wind direction could, due to its many relationships, call into question the very feasibility of the existing course of action (i.e., the firefighting plan). Even under less critical conditions it is not uncommon for the solution objectives to change several times during the decision-making process. This fourth characteristic

of complex problem situations is of particular interest. It exemplifies the tight coupling that can exist among certain problem issues, and the degree to which decision-makers must be willing to accommodate fundamental changes in the information that drives the problem situation.

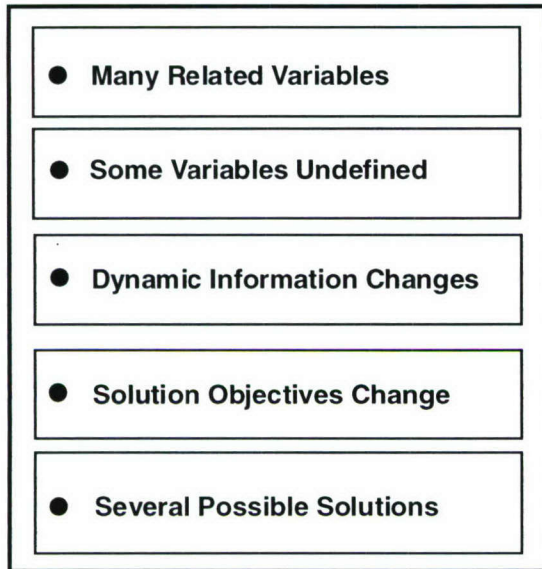


Figure 3: Character of complex problems

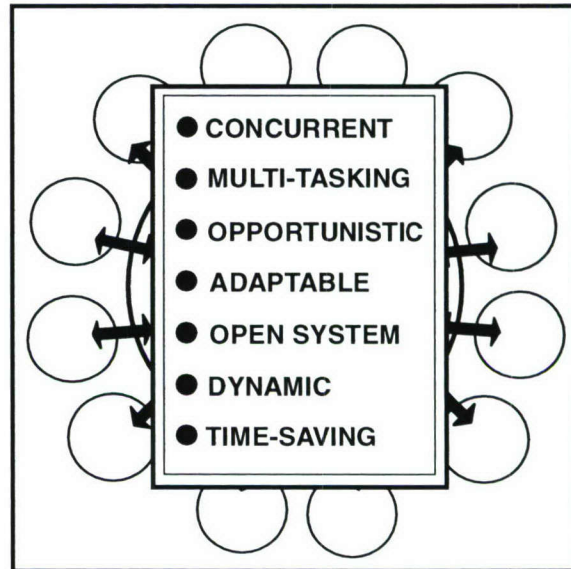


Figure 4: Parallel decision-support

Fifth, complex problems typically have more than one solution (Archea 1987). It is usually fruitless to look for an optimum solution, because there are no static benchmarks available for evaluating optimality. A solution is found to be acceptable if it satisfies certain performance requirements and if it has been determined that the search for alternatives is no longer warranted. Such a determination is often the result of resource constraints (e.g., availability of time, penalty of non-action, or financial resources) rather than a high level of satisfaction with the quality of the proposed solution.

While human decision-making in complex problem situations has so far defied rigorous scientific explanation, we do have knowledge of at least some of the characteristics of the decision-making activity.

- ◆ Decision-makers typically define the problem situation in terms of issues that are known to impact the desired outcome. The relative importance of these issues and their relationships to each other change dynamically during the decision-making process. So also do the boundaries of the problem space and the goals and objectives of the desired outcome. In other words, under these circumstances decision-making is an altogether dynamic process in which both the rules that govern the process and the required properties of the end-result are subject to continuous review, refinement and amendment.
- ◆ The complexity of the decision-making activity does not appear to be due to a high level of difficulty in any one area but the multiple relationships that exist among the many issues that impact the desired outcome. Since a decision in one area will tend to influence several other areas there is a need to consider many factors at the same time. This places a severe burden on the human cognitive system. Although the neurological mechanisms that support conscious thought

processes are massively parallel, the operation of these reasoning capabilities is largely sequential. Accordingly, decision-makers tend to apply simplification strategies for reducing the complexity of the problem-solving activity. In this regard it becomes readily apparent why 2nd Wave software provides a much more useful architecture for decision-support systems (Figure 4).

- ◆ Observation of decision-makers in action has drawn attention to the important role played by experience gained in past similar situations, knowledge acquired in the general course of decision-making practice, and expertise contributed by persons who have detailed specialist knowledge in particular problem areas. The dominant emphasis on experience is confirmation of another fundamental aspect of the decision-making activity. Problem-solvers seldom start from first principles. In most cases, the decision-maker builds on existing solutions from previous situations that are in some way related to the problem under consideration. From this viewpoint, the decision-making activity involves the modification, refinement, enhancement and combination of existing solutions into a new hybrid solution that satisfies the requirements of the given problem system. In other words, problem-solving can be described as a process in which relevant elements of past prototype solution models are progressively and collectively molded into a new solution model. Very seldom are new prototype solutions created that do not lean heavily on past prototypes.
- ◆ Finally, there is a distinctly irrational aspect to decision-making in complex problem situations. Donald Schon refers to a "...reflective conversation with the situation...". (Schon 1983). He argues that decision-makers frequently make value judgments for which they cannot rationally account. Yet, these intuitive judgments often result in conclusions that lead to superior solutions. It would appear that such intuitive capabilities are based on a conceptual understanding of the situation, which allows the problem solver to make knowledge associations at a highly abstract level.

Based on these characteristics the solution of complex problems can be categorized as an information intensive activity that depends for its success largely on the availability of information resources and, in particular, the experience and reasoning skills of the decision-makers. It follows that the quality of the solutions will vary significantly as a function of the problem-solving skills, knowledge, and information resources that can be brought to bear on the solution process. This clearly presents an opportunity for the useful employment of computer-based decision-support systems in which the capabilities of the human decision-maker are complemented with knowledge bases, expert agents, and self-activating conflict identification and monitoring capabilities.

### Principal Elements of Decision-Making

Over the past two decades that the CADRC Center has been developing distributed, collaborative decision-support systems some insights have been gained into the nature of the decision-making activity. In particular, we have found it useful to characterize decision-making in terms of six functional elements (Figure 5): *information*; *representation*; *visualization*; *communication*; *reasoning*; and, *intuition*.

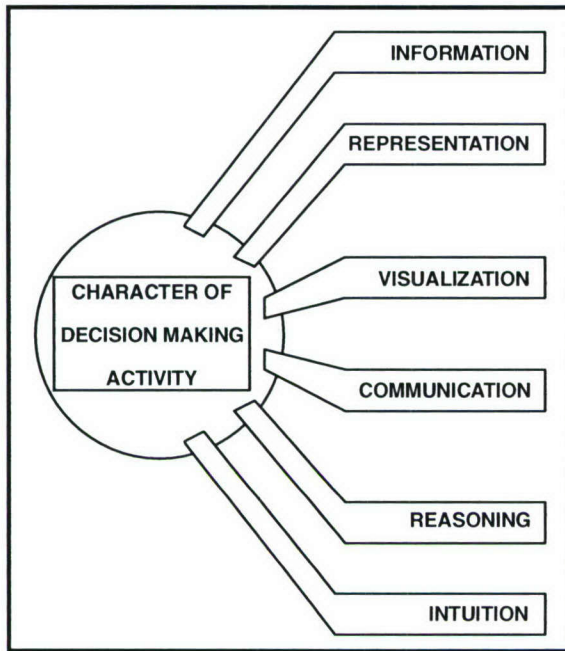
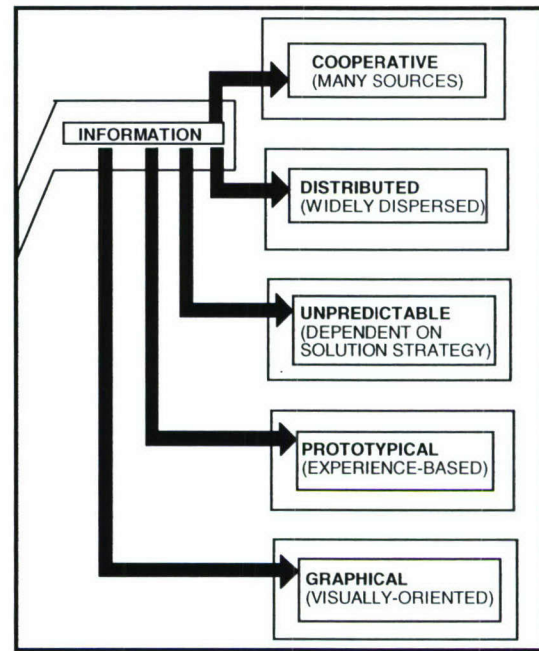


Figure 5: Decision-making elements

Figure 6: The *information* element

### The Information Element

Decision-making in complex problem situations is a collaborative activity involving many sources of information that are often widely dispersed. Seldom is all of the information required for the solution, or even only a component of the problem, physically located in the immediate vicinity of the decision-maker. In fact, much of the information is likely to reside in remote repositories that can be accessed only through electronic means, the telephone, e-mail, or the temporary relocation of a member of the problem-solving team (Figure 6). If the desired information requires expert advice the services of a consultant may be required in addition to, or instead of, access to an information resource.

The term *information* is used here in the broadest sense to include not only factual data and the progressively more comprehensive and detailed description of the problem system, but also the many knowledge bases that are part of the local and global environment within which the problem situation is constituted. In this regard, we are concerned with the knowledge of the individual members of the problem-solving team, the knowledge of peripheral players (e.g., colleagues, associates and consultants), the collective knowledge of the profession (such as the various engineering professions, the military establishment, or the management profession) and industry, and beyond that those aspects of what might be referred to as global knowledge that impact the problem context.

Typically, the problem specifications (i.e., constraints, criteria, and objectives) evolve with the problem solution as the decision-makers interact with the problem situation. Accordingly, the information requirements of the problem solver are not predictable since the information needed to solve the problem depends largely on the solution strategy adopted (Fischer and Nakakoji 1991). In this respect problem solving is a learning process in which the decision-maker progressively develops a clearer understanding of the problem that is required to be solved.

Much of the information that decision-makers use in the development of a problem solution is gleaned from experience with past projects. In fact, it can be argued that solutions commonly evolve out of the adaptation, refinement and combination of prototypes (Gero et al. 1988). This argument suggests that the more expert human decision-makers are the more they tend to rely on prototypical information in the solution of complex problems. It would appear that the accumulation, categorization and ability to apply prototype knowledge are the fundamental requirements for a human decision-maker to reach the level of *expert* in a particular domain. Based largely on the work of Gero et al. (1988) and Rosenman and Gero (1993) the following techniques used by engineering designers to develop solutions through the manipulation of prototypes can be identified as being universally applicable to other problem domains:

- ◆ **Refinement:** The prototype can be applied after changes have been made in the values of parameter variables only (i.e., the instance of the prototype is reinterpreted within the acceptable range of the parameter variables).
- ◆ **Adaptation:** Application of the prototype requires changes in the parameters that constitute the description of the prototype instance, based on factors that are internal to the prototype (i.e., a new prototype instance is produced).
- ◆ **Combination:** Application of the prototype requires the importation of parameter variables of other prototypes, producing a new instance of a reinterpreted version of the original prototype.
- ◆ **Mutation:** Application of the prototype requires structural changes to the parameter variables, either through internal manipulations or the importation of parameter variables from external sources (i.e., either a reinterpreted version of the original prototype or a new prototype is produced).
- ◆ **Analogy:** Creation of a new prototype based on a prototype that exists in another context, but displays behavioral properties that appear to be analogous to the application context.

For application purposes in knowledge-based decision-support systems prototypes may be categorized into five main groups based on knowledge content (Schon 1988, Pohl and Myers 1994):

1. **Vertical** prototype knowledge bases that contain typical object descriptions and relationships for a complete problem situation or component thereof. Such a knowledge base may include all of the types that exist in a particular problem setting, for example: an operational template for a particular kind of humanitarian relief mission; a certain type of propulsion unit; or, a building type such as a library, sports stadium, or supermarket.
2. **Horizontal** prototype knowledge bases that contain typical solutions for sub-problems such as commercial procurement practices, construction of a temporary shelter, or techniques for repairing equipment. This kind of knowledge often applies to more than one discipline. For example, the techniques for repairing a truck apply equally to the military as they do to auto-repair shops, engineering concerns, and transportation related organizations.

3. **Domain** prototype knowledge bases that contain guidelines for developing solutions within contributing narrow domains. For example, the range of structural solutions appropriate for the construction of a suspension bridge during a military mission is greatly influenced by the availability of material, the prevailing wind conditions, and the time available for erection. Posed with this design problem military engineers will immediately draw upon a set of rules that guide the design activity.
4. **Exemplar** prototype knowledge bases that describe a specific instance of an object type or solution to a sub-problem. Exemplary prototypes can be instances of vertical or horizontal prototypes, such as a particular building type or a method of welding a certain kind of steel joint that is applied across several disciplines and industries (e.g., building industry and automobile industry). Decision-makers often refer to exemplary prototypes in exploring solution alternatives to sub-problems.
5. **Experiential** knowledge bases that represent the factual prescriptions, strategies and solution conventions employed by the decision-maker in solving similar kinds of problem situations. Such knowledge bases are typically rich in methods and procedures. For example, a particularly memorable experience such as the deciding event in a past business negotiation or the experience of seeing for the first time the magnificent sail-like concrete shell walls of the Sydney Opera House, may provide the basis for a solution method that is applied later to create a similar experience in a new problem situation that may be quite different in most other respects. In other words, experiential prototypes are not bound to a specific type of problem situation. Instead, they represent techniques and methods that can be reproduced in various contexts with similar results. Experiential knowledge is often applied in very subtle ways to guide the solution of sub-problems (e.g., a subterfuge in business merger or take-over negotiations that is designed to mislead a competing party).

The amount of prototypical information is potentially overwhelming. However, the more astute and experienced decision-maker will insist on taking time to assimilate as much information as possible into the problem setting before committing to a solution theme. There is a fear that early committal to a particular solution concept might overlook characteristics of the problem situation that could gain in importance in later stages, when the solution has become too rigid to adapt to desirable changes. This reluctance to come to closure places a major information management burden on the problem solver. Much of the information cannot be specifically structured and prepared for ready access, because the needs of the problem solver cannot be fully anticipated. Every step toward a solution generates new problems and information needs (Simon 1981).

### **The Representation Element**

The methods and procedures that decision-makers utilize to solve complex problems rely heavily on their ability to identify, understand and manipulate objects (Figure 7). In this respect, objects are complex symbols that convey meaning by virtue of the explicit and implicit information that they encapsulate within their domain. For example, military strategists develop operational plans by reasoning about terrain, weather conditions, enemy positions, weapon assets, and so on. Each of these objects encapsulates knowledge about its own nature, its relationships with other objects, its behavior within a given environment, what it requires to meet its own performance objectives, and how it might be manipulated by the decision-maker within a given problem

scenario (Figure 8). This knowledge is contained in the various representational forms of the object as factual data, relationships, algorithms, rules, exemplar solutions, and prototypes.

The reliance on object representations in reasoning endeavors is deeply rooted in the innately associative nature of the human cognitive system. Information is stored in long-term memory through an indexing system that relies heavily on the forging of association paths. These paths relate not only information that collectively describes the meaning of symbols such as building, car, chair, and tree, but also connect one symbol to another. The symbols themselves are not restricted to the representation of physical objects, but also serve as concept builders. They provide a means for grouping and associating large bodies of information under a single conceptual metaphor. In fact, Lakoff and Johnson (1980) argue that “...our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature.” They refer to the influence of various types of metaphorical concepts, such as “...desirable is up” (i.e., spatial metaphors) and “...fight inflation” (i.e., ontological or human experience metaphors), as the way human beings select and communicate strategies for dealing with everyday events.

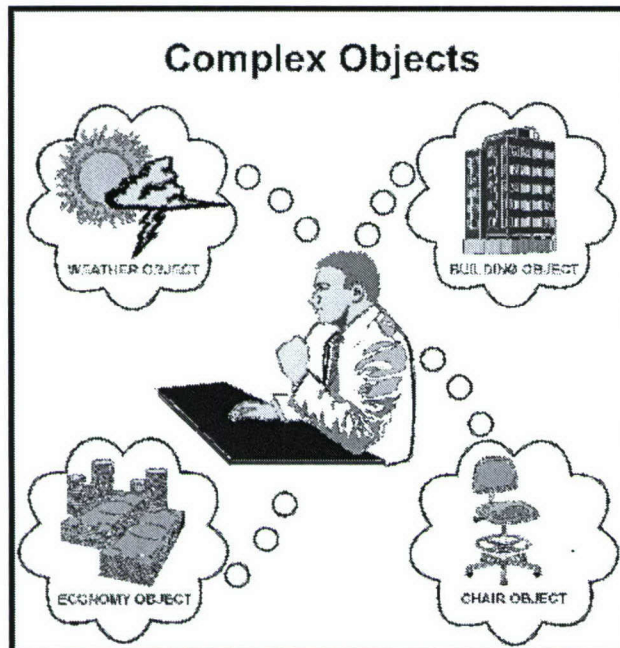


Figure 7: Symbolic reasoning with objects

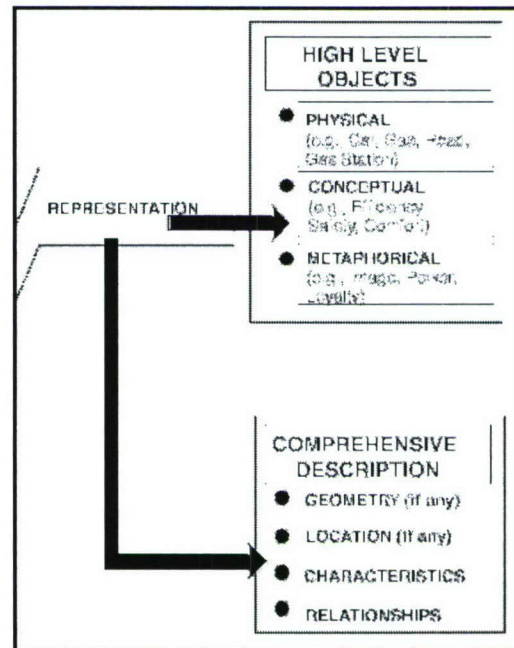


Figure 8: The *representation* element

Problem-solvers typically intertwine the factually based aspects of objects with the less precise, but implicitly richer language of metaphorical concepts. This leads to the spontaneous linkage of essentially different objects through the process of analogy. In other words, the decision-maker recognizes similarities between two or more sub-components of apparently unrelated objects and embarks upon an exploration of the discovered object seeking analogies where they may or may not exist. At times these seemingly frivolous pursuits lead to surprising and useful solutions of the problem at hand.

The need for a high level representation is fundamental to all computer-based decision-support systems. It is an essential prerequisite for embedding artificial intelligence in such systems, and forms the basis of any meaningful communication between user and computer. Without a high level representation facility the abilities of the computer to assist the human decision maker are

confined to the performance of menial tasks, such as the automatic retrieval and storage of data or the computation of mathematically defined quantities. While even those tasks may be highly productive they cannot support a partnership in which human users and computer-based systems collaborate in a meaningful and intelligent manner in the solution of complex problems.

The term *high level representation* refers to the ability of computer software to process and interpret changes in data within an appropriate context. It is fundamental to the distinction between data-centric and information-centric software. Strictly speaking data are numbers and words without relationships<sup>1</sup>. Software that incorporates an internal representation of data only is often referred to as *data-centric* software. Although the data may be represented as objects the absence of relationships to define the functional purpose of the data inhibits the inclusion of meaningful and reliable automatic reasoning capabilities. Data-centric software, therefore, must largely rely on predefined solutions to predetermined problems, and has little (if any) scope for adapting to real world problems in near real-time.

Information, on the other hand, refers to the combination of data with relationships to provide adequate context for the interpretation of the data. The richer the relationships, the greater the context and the more opportunity for automatic reasoning by software agents. Software that incorporates an internal information model (i.e., ontology) consisting of objects, their characteristics, and the relationships among those objects is often referred to as *information-centric* software. The information model provides a virtual representation of the real world domain under consideration. Since information-centric software has some *understanding* of what it is processing it normally contains tools rather than predefined solutions to predetermined problems. These software tools are commonly referred to as agents that collaborate with each other and the human user(s) to develop solutions to problems in near real-time, as they occur.

### **The Visualization Element**

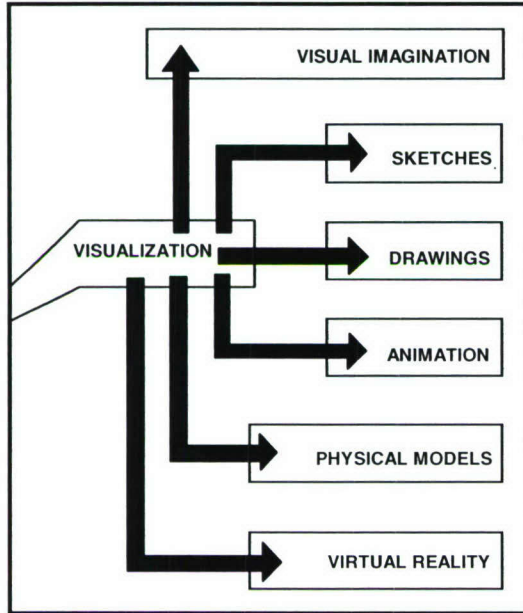
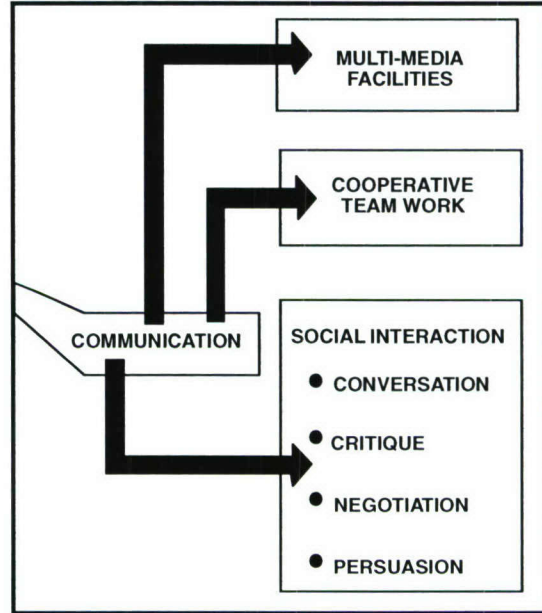
Problem solvers use various visualization media, such as visual imagination, drawings and physical models, to communicate the current state of the evolving solution to themselves and to others (Figure 9). Drawings, in particular, have become intrinsically associated with problem solving. Although the decision-maker can reason about complex problems solely through mental processes, drawings and related physical images are useful and convenient for extending those processes. The failings of the drawing as a vehicle for communicating the full intent of the decision-maker do not apply to the creator of the drawing. To the latter the drawing serves not only as an extension of long-term memory, but also as a visual bridge to its associative indexing structure. In this way, every meaningful part of the drawing is linked to related data and deliberation sequences that together provide an effectively integrated and comprehensive representation of the artifact.

From a technical point of view a great deal of headway has been made over the past two decades in the area of computer-based visualization. However, without high-level representation capabilities even the most sophisticated computer generated images are nothing but hollow shells. If the computer system does not have even the simplest understanding of the nature of the

---

<sup>1</sup> Even though data are often stored in a relational database management system, the relationships that are stored with the data in such a database are structural in nature and do not provide any information on how the data will be used (i.e., the *context* of the data).

objects that are contained in the image then it cannot contribute in any way to the analysis of those objects. On the other hand, visualization in combination with high-level representation becomes the most powerful element of the user-interface of a decision-support system. Under these circumstances, visualization promotes the required level of understanding between the user and the computer as they collaborate in the solution of a problem.

Figure 9: The *visualization* elementFigure 10: The *communication* element

### The Communication Element

The solution of complex problems is typically undertaken by a team of decision-makers. Each team member contributes within a collaborative decision-making environment that relies heavily on the normal modes of social interaction, such as conversation, critique, negotiation, and persuasion (Figure 10). Two aspects of such an interactive environment are particularly well catered for in computer-based systems. The first aspect relates to the ability of computer-driven communication networks to link together electronically based resources located anywhere on Earth or in space. Technical advances in the communication industry have greatly enhanced the ability of individuals to gain access to remotely distributed information sources, and to interact with each other over vast distances. In fact, connectivity rather than geographical distance has become the principal determinant of communication.

The second aspect is interwoven with the first by relatively recent technological advances that have permitted all types of information to be converted into digital form. Through the use of digital switching facilities modern communication networks are able to transmit telephone conversations and graphical images in the same way as data streams have been sent from one computer to another over the past 40 years.

As a direct result of these advances in communication systems the convenient and timely interaction of all of the members of a widely dispersed problem-solving team is technically assured. It is now incumbent on software developers to produce computer-based decision-support systems that can fully support collaborative teamwork, which is neither geographically

nor operationally limited. Such systems will integrate not only computer-based information resources and software agents, but also multiple human agents (i.e., the users) who will collaborate with the computer-based resources in a near real-time interactive environment.

### The Reasoning Element

Reasoning is central to any decision-making activity. It is the ability to draw deductions and inferences from information within a problem-solving context. The ability of the problem solver to reason effectively depends as much on the availability of information, as it does on an appropriately high level form of object representation (Figure 11). Decision-makers typically define complex problems in terms of issues that are known to impact the desired outcome. The relative importance of these issues and their relationships to each other change dynamically during the decision-making process. So also do the boundaries of the problem space and the goals and objectives of the desired outcome. In other words, the solution of complex problems is an altogether dynamic process in which both the rules that govern the process and the required properties of the end-result are subject to continuous review, refinement and amendment (Reitman 1964 and 1965, Rittel and Weber 1984).

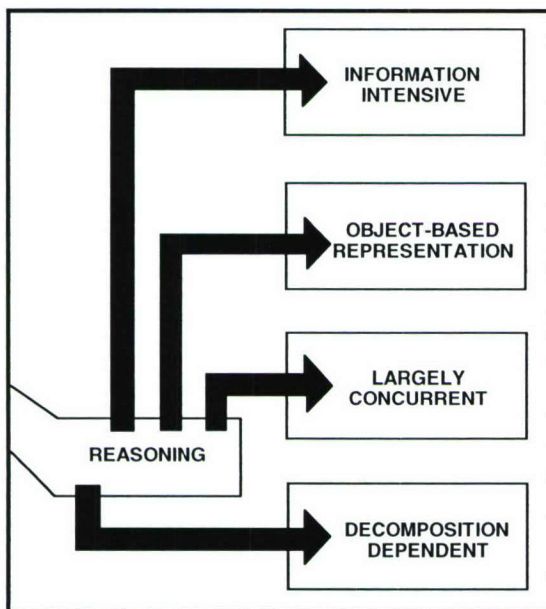


Figure 11: The *reasoning* element

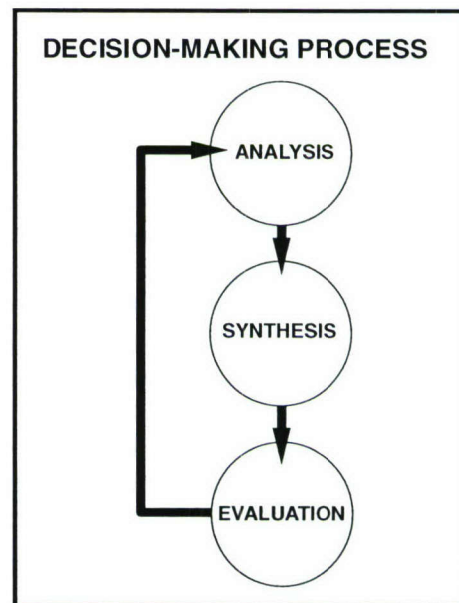


Figure 12: Reasoning methodology

As discussed previously, the complexity of a problem is normally not due to a high degree of difficulty in any one area but the multiple relationships that exist among the many issues that impact the desired outcome. Since a decision in one area will tend to influence several other areas there is a critical need for concurrency. However, the reasoning capabilities of the human problem solver are sequential in nature<sup>2</sup>. Accordingly, decision-makers find it exceedingly difficult to consider more than three or four issues at any one time. In an attempt to deal with the

<sup>2</sup> Reasoning is a logical process that proceeds in a step-by-step manner. In this respect reasoning is quite different from intuition, which allows humans to spontaneously come to conclusions that are neither consciously formulated nor explainable at the time of their first appearance.

concurrency requirement several strategies are commonly employed to reduce the complexity of the reasoning process to a manageable level.

- ◆ **Constraint Identification:** By sifting through the available information the problem-solver hopes to find overriding restrictions and limitations that will eliminate knowledge areas from immediate consideration.
- ◆ **Decision Factor Weighting:** By comparing and evaluating important problem issues in logical groupings, relative to a set of predetermined solution objectives, the decision-maker hopes to identify a smaller number of issues or factors that appear to have greater impact on the final solution. Again, the strategy is to reduce the size of the information base by early elimination of apparently less important considerations.
- ◆ **Solution Conceptualization:** By adopting early in the decision-making process a conceptual solution, the problem-solver is able to pursue a selective evaluation of the available information. Typically, the problem-solver proceeds to subdivide the decision factors into two groups, those that are compatible with the conceptual solution and those that are in conflict. By a process of trial and error, often at a superficial level, the problem-solver develops, adapts, modifies, re-conceives, rejects and, often, forces the preconceived concept into a final solution.

In complex problem situations reasoning proceeds in an iterative fashion through a cycle of *analysis*, *synthesis* and *evaluation* (Figure 12). During the *analysis* stage (Figure 13) the problem-solver interprets and categorizes information to establish the relative importance of issues and to identify compatibilities and incompatibilities among the factors that drive these issues.

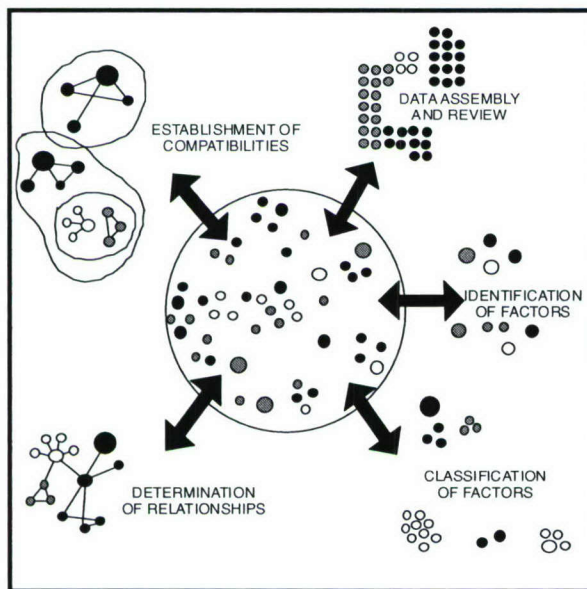


Figure 13: *Analysis* stage of reasoning

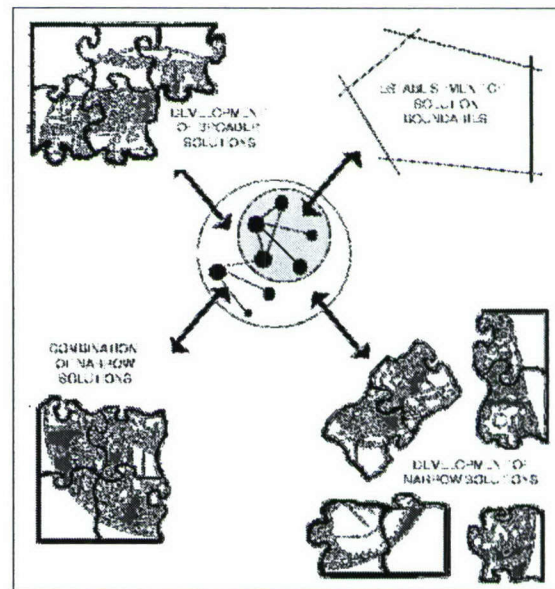


Figure 14: *Synthesis* stage of reasoning

During *synthesis* (Figure 14) solution boundaries and objectives are continuously reexamined as the decision-maker develops narrow solutions to sub-problems and combines these narrow solutions into broader solutions. Initially, these solution attempts are nothing more than trial

balloons. Or, stated in more technical terms, explorations based on the development of the relationships among the principal issues and compatible factors identified during the *analysis* stage. Later, as the problem-solving activity progresses, firmer conceptual solution strategies with broader implications emerge. However, even during later cycles the solution strategies tend to be based on a limited number of issues or factors.

During the *evaluation* stage (Figure 15) the decision-makers are forced to test the current solution strategy with all of the known problem issues, some of which may have been considered only superficially or not at all during the formulation of the current solution proposal. This may require the current solution concepts to be modified, extended or altogether replaced. Typically, several solution strategies are possible and none are completely satisfactory. Archea (1987), in his description of the architectural design activity refers to this activity as "... *puzzle-making*", suggesting by implication that the decision-maker utilizes the reasoning cycle more as a method for exploring the problem space than as a decision-making tool for forcing an early solution.

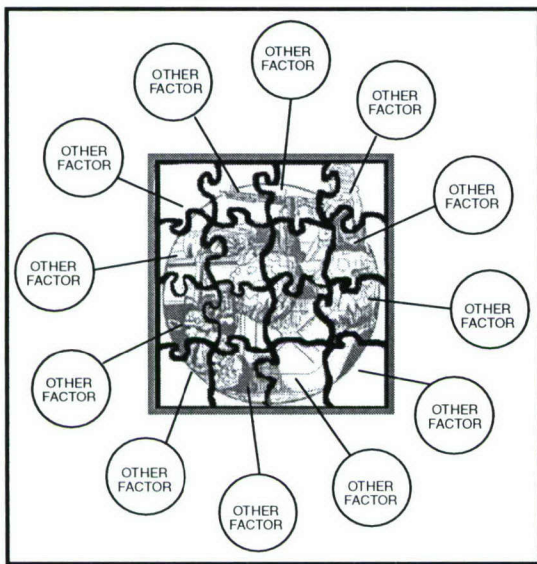


Figure 15: *Evaluation* stage of reasoning

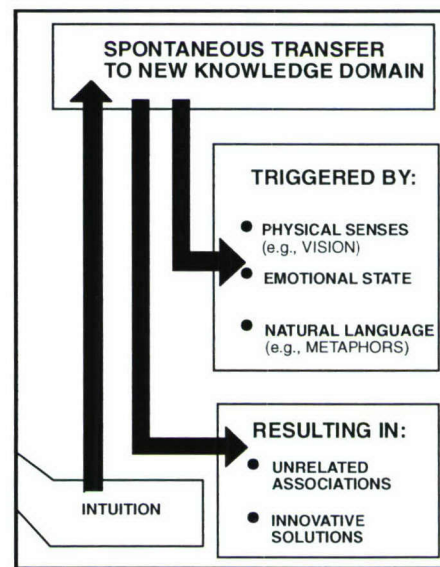


Figure 16: The *intuition* element

### The Intuition Element

Donald Schon (1983 and 1988) has written extensively about the intuitive aspects of decision-making. Although he focused primarily on engineering design as an application area, his views provide valuable insight into the solution of complex problems in general. Design has all of the common characteristics of complex problem situations, and some additional ones such as the desire for solution uniqueness, that make it a prime candidate for computer-based assistance (Pohl et al.1994).

In Schon's (1988) view designers enter into "... *design worlds*" in which they find the objects, rules and prototype knowledge that they apply to the design problem under consideration. The implication is that the designer continuously moves in and out of design worlds that are triggered by internal and external stimuli. While the reasoning process employed by the designer in any particular design world is typically sequential and explicitly logical, the transitions from state to state are governed by deeper physiological and psychological causes. Some of these causes can

be explained in terms of associations that the designer perceives between an aspect or element of the current state of the design solution and prototype knowledge that the designer has accumulated through experience. Others may be related to emotional states or environmental stimuli, or interactions of both (Figure 16).

For example, applying Schon's view to the broader area of complex problem solving, a particular aspect of a problem situation may lead to associations in the decision-maker's mind that are logically unrelated to the problem under consideration. However, when the decision-maker pursues and further develops these associations they sometimes lead to unexpected solutions. Typically, the validity of these solutions becomes apparent only after the fact and not while they are being developed. In popular terms we often refer to these solutions as *creative leaps* and label the author as a brilliant strategist. What we easily forget is that many of these intuitions remain unrelated associations and do not lead to any worthwhile result. Nevertheless, the intuitive aspect of decision-making is most important. Even if only a very small percentage of these intuitive associations were to lead to a useful solution, they would still constitute one of the most highly valued decision-making resources.

The reasons for this are twofold. First, the time at which the decision-maker is most willing to entertain intuitive associations normally coincides with a most difficult stage in the problem solving process. Typically, it occurs when an impasse has been reached and no acceptable solution strategy can be found. Under these circumstances intuition may be the only remaining course of action open to the decision-maker. The second reason is particularly relevant if there is a strong competitive element present in the problem situation. For example, during a chess game or during the execution of military operations. Under these circumstances, strategies and solutions triggered by intuitive associations will inevitably introduce an element of surprise that is likely to disadvantage the adversary.

The importance of the *intuition* element itself in decision-making would be sufficient reason to insist on the inclusion of the human decision-maker as an active participant in any computer-based decision system. In designing and developing such systems in the CADRC over the past decade we have come to appreciate the importance of the human-computer partnership concept, as opposed to automation. Whereas in some of our early systems (e.g., ICADS (Pohl et al. 1988) and AEDOT (Pohl et al. 1992)) we included agents that automatically resolve conflicts, today we are increasingly moving away from automatic conflict resolution to conflict detection and explanation. We believe that even apparently mundane conflict situations should be brought to the attention of the human agent. Although the latter may do nothing more than agree with the solution proposed by the computer-based agents, he or she has the opportunity to bring other knowledge to bear on the situation and thereby influence the final determination.

### **The Human-Computer Partnership**

To look upon decision-support systems as partnerships between users and computers, in preference to automation, appears to be a sound approach for at least two reasons. First, the ability of the computer-based components to interact with the user overcomes many of the difficulties, such as representation and the validation of knowledge, that continue to plague the field of machine learning (Forsyth 1989, Thornton 1992, Johnson-Laird 1993). Second, human and computer capabilities are in many respects complementary (Figures 17 and 18). Human capabilities are particularly strong in areas such as communication, symbolic reasoning,

conceptualization, learning, and intuition. We are able to store and adapt experience and quickly grasp the overall picture of even fairly chaotic situations. Our ability to match patterns is applicable not only to visual stimuli but also to abstract concepts and intuitive notions. However, although the biological bases of our cognitive abilities are massively parallel, our conscious reasoning capabilities are essentially sequential. Therefore, large volumes of information and multi-faceted decision contexts tend to easily overwhelm human decision-makers.

When such an overload occurs we tend to switch from an analysis mode to an intuitive mode in which we have to rely almost completely on our ability to develop situation awareness through abstraction and conceptualization. While this is our greatest strength it is also potentially our greatest weakness. At this intuitive meta-level we become increasingly vulnerable to emotional influences that are an intrinsic part of our human nature and therefore largely beyond our control.

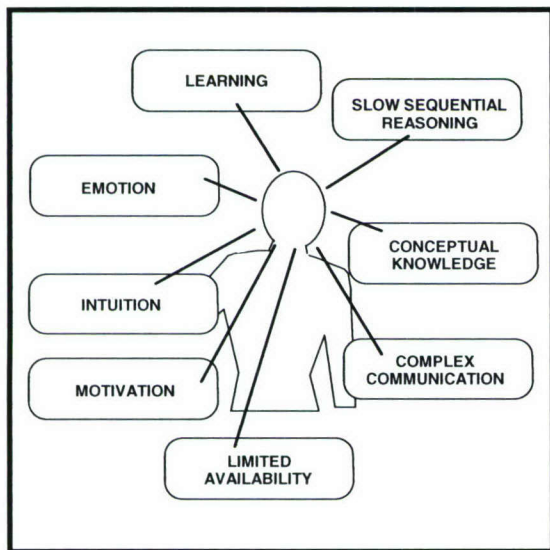


Figure 17: Human abilities and limitations

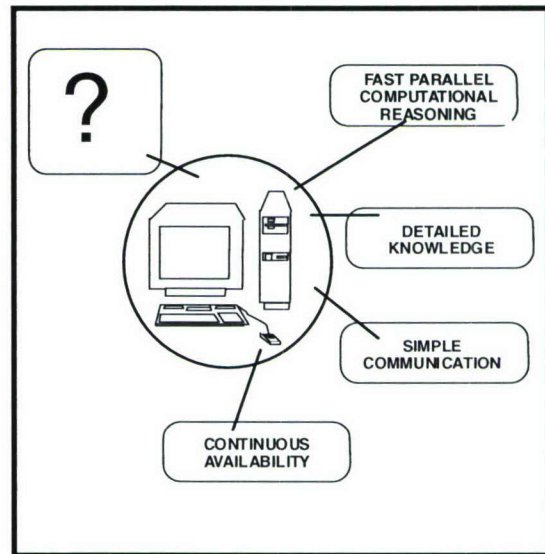


Figure 18: Computer abilities and limitations

The capabilities of the computer are strongest in the areas of parallelism, speed and accuracy (Figure 18). Whereas the human being tends to limit the amount of detailed knowledge by continuously abstracting information to a higher level of understanding, the computer excels in its almost unlimited capacity for storing data. While the human being is prone to making minor mistakes in arithmetic and reading, the computer is always accurate. A slight diversion may be sufficient to disrupt our attention to the degree that we incorrectly add or subtract two numbers. However, if the error is large we are likely to notice that something is wrong further downstream due to our ability to apply conceptual checks and balances. The computer, on the other hand, cannot of its own accord (i.e., at the hardware level) distinguish between a minor mistake and a major error. Both are a malfunction of the entirely predictable behavior of its electronic components. However, at the software level it is possible to provide a layer of automatic reasoning capabilities (i.e., collaborating agents) served by an underlying information model (i.e., ontology). Software with such embedded capabilities is able to draw inferences leading to more sophisticated human-like conclusions.

The differences between the human being and the computer are fundamental. All of the capabilities of the digital computer are derived from the simple building blocks of 0 and 1. There is no degree of vagueness here, 0 and 1 are precise digital entities and very different from the

massively parallel and largely unpredictable interactions of neurons and synapses that drive human behavior. It is not intuitively obvious how to create the high level representations of real world objects (e.g., ship, aircraft, dog, house, power, security, etc.) that appear to be a prerequisite for reasoning and learning capabilities, in a digital computer. While these objects can be fairly easily represented in the computer as superficial visual images (in the case of physical objects such as aircraft, weapons and buildings) and data relationships (in the case of conceptual objects such as power and security) that in itself does not ensure that the computer has any understanding of their real world meaning. These representations are simply combinations of the basic digital building blocks that model, at best, the external shell rather than the internal meaning of the object.

In this respect the term *information-centric* refers to the representation of information in the computer, not to the way it is actually stored in a digital machine. This distinction between *representation* and *storage* is important, and relevant far beyond the realm of computers. When we write a note with a pencil on a sheet of paper, the content (i.e., meaning) of the note is unrelated to the storage device. A sheet of paper is designed to be a very efficient storage medium that can be easily stacked in sets of hundreds, filed in folders, bound into volumes, folded, and so on. However, all of this is unrelated to the content of the written note on the paper. This content represents the meaning of the sheet of paper. It constitutes the purpose of the paper and governs what we do with the sheet of paper (i.e., its use). In other words, the nature and efficiency of the storage medium is more often than not unrelated to the content or representation that is stored in the medium.

In the same sense, the way in which we store bits (i.e., *0s* and *1s*) in a digital computer is unrelated to the meaning of what we have stored. For a computer to interpret data it requires an information structure that provides at least some level of *context*. This can be accomplished utilizing an ontology of objects with characteristics and a rich set of relationships to create a virtual version of a real world situation. The resultant level of information representation is normally adequate to provide the context within which agent logic can automatically operate.

## References

- Archea J. (1987); 'Puzzle-Making: What Architects Do When No One Is Looking'; in Kalay (ed.) Principles of Computer-Aided Design: Computability of Design'; Wiley, New York, New York.
- Bohm D. (1983); 'Wholeness and the Implicate Order'; Ark Paperbacks, London, England.
- Schon D. (1983); 'The Reflective Practitioner: How Professionals Think in Action'; Basic Books.
- Fischer G. and K. Nakakoji (1991); 'Making Design Objects Relevant to the Task at Hand'; Proc. AAAI-91 Ninth National Conference on Artificial Intelligence, MIT Press, Cambridge, Massachusetts (pp.67-73).
- Forsyth R. (1989); 'Machine Learning: Principles and Techniques'; Chapman and Hall, Computing Series, London, England.
- Gero J., M. Maher and W. Zhang (1988); 'Chunking Structural Design Knowledge as Prototypes'; Working Paper, The Architectural Computing Unit, Department of Architectural and Design Science, University of Sydney, Sydney, Australia.
- Johnson-Laird P. (1993); 'Human and Machine Thinking'; Erlbaum, Hillsdale, New Jersey.

Lakoff G. and M. Johnson (1980); 'Metaphors We Live By'; University of Chicago Press, Chicago, Illinois.

Pohl J., A. Chapman, L. Chirica, R. Howell and L. Myers (1988); 'Implementation Strategies for a Prototype ICADS Working Model'; Technical Report, CADRU-02-88, CAD Research Center, Design and Construction Institute, College of Architecture and Environmental Design, Cal Poly, San Luis Obispo, California.

Pohl J., J. La Porta, K. Pohl and J. Snyder (1992); 'AEDOT Prototype (1.1): An Implementation of the ICADS Model'; Technical Report, CADRU-07-92, CAD Research Center, Design and Construction Institute, College of Architecture and Environmental Design, Cal Poly, San Luis Obispo, California.

Pohl J. and L. Myers (1994); 'A Distributed Cooperative Model for Architectural Design'; Automation in Construction, 3(2-3), July (pp.177-185).

Pohl J., L. Myers and A. Chapman (1994); 'Thoughts on the Evolution of Computer-Assisted Design'; Technical Report, CADRU-09-94, CAD Research Center, Design and Construction Institute, College of Architecture and Environmental Design, Cal Poly, San Luis Obispo, California, September.

Reitman W. (1964); 'Heuristic Decision Procedures, Open Constraints, and the Ill-Defined Problems'; in Shelley and Bryan (eds.) Human Judgments and Optimality, Wiley, New York, New York (pp.282-315).

Reitman W. (1965); 'Cognition and Thought'; Wiley, New York, New York.

Rittel H. and M. Webber (1984); 'Planning Problems are Wicked Problems'; in Cross (ed.) Developments in Design Methodology, Wiley, New York, New York (pp.135-144).

Rosenman M. and J. Gero (1993); 'Creativity in Design Using Design Prototypes', in Gero and Maher (eds.) Modeling Creativity and Knowledge-Based Creative Design, Erlbaum, Hillsdale, New Jersey (pp.111-138).

Schon D. (1988); 'Designing: Rules, Types and Worlds'; Design Studies, 9(3), July (pp.181-190).

Senge P. (1993); 'Transforming the Practice of Management'; Human Resource Development Quarterly, Jossey-Bass, San Francisco, California.

Simon H. (1981); 'The Sciences of the Artificial'; MIT Press, Cambridge, Massachusetts.

Thornton C. (1992); 'Techniques in Computational Learning'; Chapman and Hall, Computing Series, London, England.

## **A Service-Oriented Solution to Semantic-Interoperability among Expressive Systems**

**Kym Pohl**  
**CDM Technologies Inc.**  
**2975 McMillan Ave.**  
**San Luis Obispo, CA. 93401**  
**805-541-3750 x233**  
**Fax: 805-541-8296**  
[kpohl@cdmtech.com](mailto:kpohl@cdmtech.com)  
<http://www.cdmtech.com>  
<http://www.cadrc.calpoly.edu>

### **ABSTRACT**

The need for expressive, potentially disparate systems to interoperate in meaningful and useful ways is increasingly gaining focus in today's arena of information-based, decision-support systems. Unlike data-oriented exchange, such interoperability must go beyond the elementary communication of discrete values and endeavor to support a more expressive context-oriented inter-system interaction. A key issue with such functionality is the support, moreover the promotion, of meaningful interoperability in a manner that allows individual systems to still perceive a domain in accordance with their own native representation. In other words, the meaningful integration of disparate systems in a manner that allows each collaborating system to retain its potentially unique means of representing, or perceiving, the domain over which it operates. In the past, several approaches to this problem have been postulated, including the development of a dedicated translation facility for each source/target combination, development of a universal ontology integrating domains from all systems into a cohesive single representation, and so on. However, these solutions have proven problematic at best. Specific, one-off translators are usually tightly coupled with both systems and have limited maintainability leading to inadequate support for dealing with inevitable representational changes. The alternative approach of bridging systems through developing a single, universal representation is not only highly impractical but also requires an ongoing effort of monumental proportions to achieve even a remotely acceptable solution. The resulting solution typically fails to achieve the levels of expressiveness and precision needed by highly analytical decision-support systems. As such, a more suitable solution is required that not only retains individual perspective but also supports translation of semantically rich content among potentially diverse representations. Considering this inherent complexity it would appear that a suitable solution will require the employment of reasoning-enabling technologies capable of supporting the high level analysis involved in performing such context-based translation. Above and beyond the need for complex translation among differing perspectives, due to the service and web-oriented environments in which many decision-support systems operate there is an additional opportunity to fashion a solution around a Service-Oriented Architecture (SOA) compatible with a web services environment. In this paradigm, both formalized and more ad hoc system capabilities could be exposed to collaborating systems as remotely accessible services. Interoperability in this sense involves systems interacting with each other in terms of services. Reliant on support for complex translation to map between potentially expressive perspectives, this notion of remote service invocation offers a simple yet effective metaphor for addressing the increasing need for useful

interaction among potentially disparate systems. The focus of this paper is to provide a vision and supporting design for a service-oriented interoperability bridge capable of supporting both standalone and web services-based interoperability among systems operating over potentially disparate and expressive representations. Exploiting capabilities offered by both the artificial intelligence and semantic web-based worlds the presented design incorporates technologies such as inference engines, rule-based systems, XML, XSLT, web services and service-oriented architectures to provide the needed infrastructure to support meaningful interoperability among context-based systems in an information age. This paper will be presented in the invited session “Ontology Driven Interoperability for Agile Applications using Information Systems: Requirements and Applications for Agent Mediated Decision Support” of WMSCI 2006.

**Keywords:** semantic interoperability, ontology, service-oriented architecture, web services, translation, reasoning, XML, XSLT, decision-support, information systems, context

## 1. THE NEED FOR SEMANTIC INTEROPERABILITY

The topic of semantic interoperability is increasingly gaining importance in today’s information age. With the context-sensitive analytical demands placed on today’s software-based decision-support systems it is more critical than ever for interaction among such systems to go beyond the simple sharing of data and venture into the realm of semantic exchange. Context-oriented systems, as opposed to data-centric systems, rely heavily on the contextual depth of the descriptions over which they operate. Contextual depth, or semantic expression, forms the fundamental enabler for such systems to offer users meaningful assistance in the areas of resource-allocation, threat-analysis, and conflict-resolution to name a few. Whether interacting with other context-enabled systems or accepting data feeds from more data-centric systems, the need to communicate context as opposed to simply structure no doubt places a significantly higher burden on the representational depth of the overall exchange. Providing support for such context-centric interoperability is the topic of much research within academia and industry alike [15].

Perhaps not surprisingly, it is the fundamental strength of context-centric, decision-support systems that poses the most challenging problem to this endeavor. This critical enabler, and simultaneous nemesis, is *representational expressiveness*. As the term implies, the context-oriented approach to building decision-support systems endeavors to go beyond the classical *nuts and bolts* approach to representation (i.e., isolated chunks of typically numeric or string-based data with little or no inter-relationships and essentially void of any derivable meaning) and incorporate the potentially numerous relationships, constraints, and business rules that are needed for the more complex analysis inherent in agent-based, decision-support analysis. A critical aspect of such representational depth is *perspective*. The biases associated with how a particular entity or concept is *viewed* are quite significant to the decision-making process. As such, perspective is a critical ingredient to truly effective context-oriented representation. For example, perceiving a truck as a sequenced collection of assemblies may be more appropriate, and effective, to an assembly-line management decision-support system than as a means of transporting cargo, the latter perhaps being more appropriate for a logistics support system. However, in the scaled reality of today’s multi-dimensional software-dependent environment, the eventual dilemma created by each context-based system essentially *seeing the world* according to

its own personalized, albeit expressive, domain model (i.e., ontology) is the increased disparity that inevitably develops between the representations upon which these systems operate. In other words, the most empowering ingredient in context-oriented computing also presents one of the most perplexing issues when such systems are required to interoperate [18] [20].

To support meaningful collaboration with this decision-support environment there must be a translational component capable of mapping expressive context in a manner that preserves individual perspective, and does so in a fashion transparent to both parties. It is this focus that drove the design and development of the interoperability facility presented in this discussion. The following sections describe the criteria employed in developing an suitable solution to this paradigm, the various technologies exploited to realize its implementation, and finally, a discussion of how these pieces are combined to form an effective design capable of supporting expressive interoperability between context-oriented systems.

### **Criteria for an Effective Solution**

To successfully address these issues, candidate technologies must satisfy several criteria. Forefront among such criteria is adherence to standards and substantial support for flexibility and reuse.

One of the primary goals of any solution intended for repeated application to varying interoperability scenarios is the ability to be flexible. A key aspect to such adaptability is the clear separation of framework from application specifics. In other words, a candidate framework should support the various abstractions associated with translation-based interoperability among heterogeneous system representations. Some of these abstracted concepts include the basic *service* of translation regardless of the actual transformation paradigm or specific translation rules employed. This framework should also identify the necessary interfaces supporting connection to the application-specific side of the equation. The latter requiring the necessary functionality and blueprint to effectively *adapt* client systems to the particular interoperability framework and interaction model presented by the solution.

Another significant property of any candidate solution is the support and promotion of available industry standards. This is particularly significant when a high degree of reusability is intended. Accordingly, the application of such a tool should center on industry-familiar technologies, standards, and tools. Not only does adherence to available standards offer a helpful guide in developing a particular solution but also aids in avoiding significant consequences in a field where complexity and *one-off* solutions abound.

As stated in the previous section, a critical ingredient of interoperability among context-oriented systems is the preservation of individual perspective. Successfully addressing this issue requires the ability to understand the subtleties inherent in such a concept (e.g., implied domain-specific constraints that have equally obscure and individual counterparts when considered from other domain-related perspectives). The considerations involved in translating between such perspectives often require a level of reasoning approaching the realm of expert systems. In many respects, for the more complex context-oriented interaction a level of decision-support on par with solutions supporting multi-variable, complex problems is needed. In fact, in certain cases correct contextual translation may even require the inclusion of human decision makers to represent higher-level concepts and considerations not able to be sufficiently represented within the encoded translation logic itself.

An additional, yet often overlooked, criterion for a successful solution to the interoperability problem is not only the support of expressive context translation but also the ability to support a more straightforward transformation without incurring the overhead associated with support for the former. This is often the case where solutions targeting complex problems offer over-engineered and subsequently inefficient solutions for less taxing scenarios. The goal is to support a range of complexities and to limit any incurred overhead to situations where it cannot be avoided.

## 2. SELECT INGREDIENTS IN SUPPORT OF THE TARGET CRITERIA

The solution presented in this discussion takes the form of a reusable service-oriented interoperability bridge framework. As such, this facility presents its users with a well-structured set of functional modules, interfaces for adaptation to the specifics of its client systems, and an overarching programming model employing a strong *remote service request* metaphor. The orientation toward such a metaphor allows each interoperating system to view its counterparts, as well as expose itself, as a collection of available services within a virtual homogeneous representational environment. The resulting interoperability model promotes a decoupled architecture requiring no notion of system identity other than the standard service descriptors registered with the bridge's services registry. Further, due to the promotion of a virtual homogeneous representation across interoperating systems, bridge clients (i.e., service users and service providers) can effectively interact with each other in terms of their own specific native *languages*.

### Enabling Technologies

The interoperability bridge incorporates a number of prominent technologies to accomplish its objective of seamless semantic interoperability within an environment of potentially diverse and expressive representations. First among this set of technologies is an optional web services interface [2] [3] [10] [11] [12] [13]. In either standalone or this web services form, by supporting standardized service lookup registries and interaction protocols, the accompanying service-oriented architecture presents an extensible, decoupled capability-oriented model for system interaction. In this model services are employed by bridge clients on an as-needed basis allowing the classical notion of operational boundaries to effectively expand and contract as the need for various capabilities varies throughout time. Furthermore, and especially in its web services form, standard service interaction protocols are employed empowering clients with an ability for discovering (i.e., their existence, not necessarily their nature) and engaging new capabilities. Although the issue of semantic discovery is still an area of significant research, web services architectures lay an effective foundation for such discovery-oriented dynamics.

The eXtensible Markup Language (XML) [9] [14] together with its Extensible Stylesheet Language Transform (XSLT) [1] language counterpart are two additional technologies employed by the interoperability bridge. XML provides a flexible means of defining structure through the use of syntactical tags. XML schemas can be developed to describe the structural aspects of entities, notions, and concepts. Receivers can process XML documents based on these schemas in an interpretational manner. The result, although accompanied by a significant caveat, is a means whereby software components can process incoming content based on a previously unknown representation. However, it should be noted that such discovery is limited to structural characteristics and does not include the discovery of semantics, *or meaning*, vital in context-

oriented decision-support systems. Even considering a schema describing the domain of concepts, rules, and implications, there is still the requirement for processors of such content to be equipped with a pre-conceived understanding of the semantics of even this very abstract domain. At some point, the semantics need to be adequately represented beyond simple structure. That said, however, structural detection does play a significant role in the eventual goal of true contextual discovery but should still be viewed as only one piece of the puzzle.

Perhaps one of the more significant additions to the world of tag-based schema languages is the ability to formally describe inter-schema transformation rules in an XML-friendly syntax. XSLT is a language that can be used to describe exactly how content based on one XML schema can be mapped into another. Whether defined statically or dynamically at runtime, XSL transforms can be effectively applied for straightforward, property-to-property translations.

Although translation at this level is useful, for the more complex transformation inherent in context-oriented representations a more powerful paradigm is required. The additional reasoning required for this level of transformation can be successfully addressed through the use of inference engine-based technology. Similar to XSLT, inference engine-based translation represents transformation logic as sets of managed rules. However, as opposed to the aforementioned XSLT mappings, these rule sets can embody significantly more complex logic supporting extensive condition-based pattern matching and the subsequent management of progressively satisfied, or partial, pattern matches. Some examples of rule-based inference engines include the CLIPS expert system shell developed by NASA [17] and a somewhat similar, although more easily embeddable within Java-based environments, JESS inference engine developed by Sandia Laboratories [6]. In either case, complex transformation logic can be implemented as pseudo expert systems applying various levels of reasoning to determine the appropriate transformation given an expressive contextual fragment. An illustration of the benefits of such capabilities would be the case, for example, where the transformation of a heavily constrained plan requiring its *truth* maintained as it moves from one representational world to another. Under these circumstances, assured truth maintenance may require a level of decision-support capable of reorganizing the plan in a manner that complies with the additional constraints described in the target *world* while still representing the initial intent, or goals, outlined in the source *world*. Availability of this level of transformation capability allows such translation-oriented decision-making to be both functionally and architecturally encapsulated within the bridge itself, transparent to its clients. Given sufficient mapping logic by which to configure this reasoning environment, content passing between bridge clients can be made fully compliant with the various native constraints. The resulting interoperability model avoids the representational contamination associated with requiring clients to distinguish foreign content from their own native content requiring a level of additional local processing to become compliant with local constraints before it can be reacted to in native terms.

### 3. A SOLUTION: THE INTEROPERABILITY BRIDGE

The interoperability bridge architecture is essentially comprised of three components. These include the Service Center, an expandable pool of pluggable translators, and individual client Connectors. The following section discusses each of these components in detail.

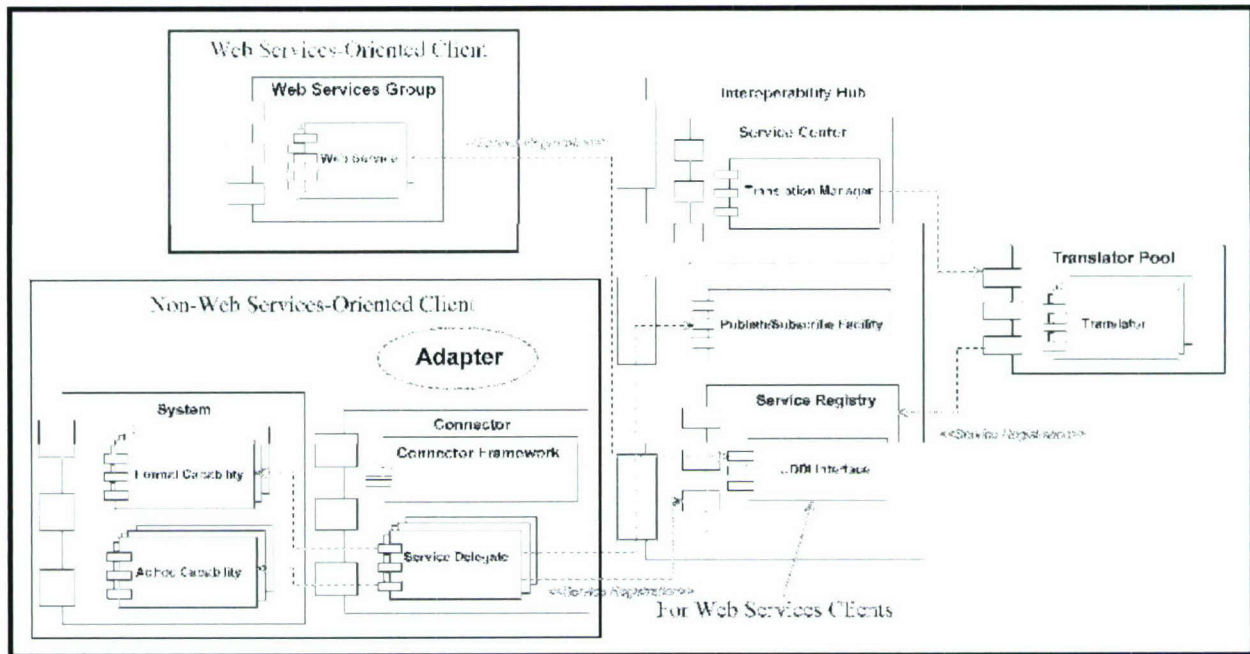


Figure 1 – The Interoperability Bridge, Illustrated Here in its Web Services-Enabled Form, Promotes Meaningful Interoperability Between Collections of Both Formal and Ad Hoc Capabilities

## The Service Center

As its name implies, the Service Center provides the high level management required to accept and process client requests. As the main interface for bridge clients this component is exposed in either a standalone or web services form. Exposed as a web service, this component adheres to a standard web services framework [10]. Configured in either manner, the form in which this component is exposed has no affect on the remainder of components forming the remainder of the bridge. Figure 1 provides an illustration of how the Service Center promotes a service-oriented architecture consisting of both formal and ad hoc capabilities. In its web services form, the Service Center provides a web-enabled facility where prospective clients can discover registered services using Universal Description, Discovery, and Integration (UDDI)-based registry lookup and invoke such services via XML structured content over SOAP-based (Simple Object Access Protocol) [21] communication. As is illustrated in Figure 1, services can exist as fully formalized services or in a more ad hoc manner that can be adapted to this service-oriented paradigm through employment of the Connector Framework (*the Connector Framework will be discussed in greater depth in a later section*). Regardless of the formality and sophistication of services, the Service Center presents its clients with an aggregated and cohesive view of exposed capabilities offered by its currently connected clientele and does so in a manner irrespective of their native representation.

Apart from providing the main client interface for the bridge, the second duty of the Service Center is to effectively broker inter-client communication to the appropriate member of the Translator Pool (*discussed in the next section*). The Service Center manages this activity in the form of a Translation Manager. Not a translator itself but rather a manager of such facilities, the

Translation Manager is responsible for discovering and engaging suitable translation services to perform the required representational mappings. The Translation Manager is capable of interfacing with candidate translators in either a standalone or web services form.

### **The Translator Pool**

Collectively known as the Translator Pool, and the second of the three primary components comprising the bridge architecture, the particular set of translators registered with the bridge determine the breadth and depth of interoperability available to bridge clients at any given time. As mentioned earlier, these translators interface with the bridge in terms of a *service plug-in* architectural pattern. In other words, various translators, some of which may be developed by third-parties as long as they support the specified translator interface, register with the main bridge manager effectively exposing their particular flavor of translation to prospective bridge clients. Whether performed in a web services manner or as a standalone capability, registration takes the form of conveying both the source and target XML Schema Definition (XSD) namespace pair that the particular translator is capable of mapping between. The specific translation paradigm employed by a particular translator is irrelevant outside the scope of the translator itself and is therefore completely shielded from bridge clients as well as the other bridge components (*more will be discussed on translation paradigms in the next section*). Once registered, the Translation Manager routes inter-client content to the appropriate translator within the pool in terms of a standard brokering model. Once invoked, the particular translator applies its preconfigured mapping rules to perform the specific mapping operation. Once translation is complete the results are sent to those bridge clients that have registered interests with the bridge for such translation events. In this manner, interaction between a translator and receiving clients is effectively decoupled to the point of a common publish/subscribe facility internal to the bridge.

### **The Client Connector**

The third component to this interoperability solution is the Client Connector. Each client wishing to invoke the services offered through the bridge (i.e., publication, query, subscription, or any conceivable capability exposed by a remote client) does so via its dedicated client connector. Based on a reusable framework provided as part of the bridge, a connector is built for each type of bridge client in accordance with the client's particular native interface. In essence, these connectors effectively *adapt* a client to seamlessly operate within the interoperability bridge's interaction model. This approach has a very significant advantage over more intrusive connection models. Essentially acting as a proxy for its specific client, the connector front-end effectively isolates native client functionality from any preconceived knowledge of the bridge or any remote clients it may be indirectly connected to. Incoming service requests are processed into native calls by the receiving client's connector. For outgoing requests, clients have a choice as to the level of knowledge it has regarding the existence of a remote *world*. For client's wishing to initiate remote requests, such invocation logic may be woven into the functional fabric of the client, or, may be solely isolated within the connector implementation itself. The latter of these two configuration extremes exploits the *delegate* design pattern implemented within the connector architecture itself. For clients unaware of the existence of an inter-client bridge, such delegates can react to locally occurring events by themselves issuing remote service requests against the bridge. In either case, the results of inter-client service invocations can be asynchronously returned as requests to the issuer's publication service, in whatever form that capability may be.

## Anatomy of the Client Connector Framework

To assist in coping with differing flavors of client interfaces, the Interoperability Bridge platform offers a framework for building and managing client connectors. This framework consists of well-defined interfaces, functional modules, and a blueprint (i.e., design pattern) for combining these components with client-specific behavior. Figure 2 illustrates the architecture of the Connector Framework along with the various interfaces on which client-specific concerns are implemented. The following section describes the workflow across the various connector components for both outgoing and incoming communications.

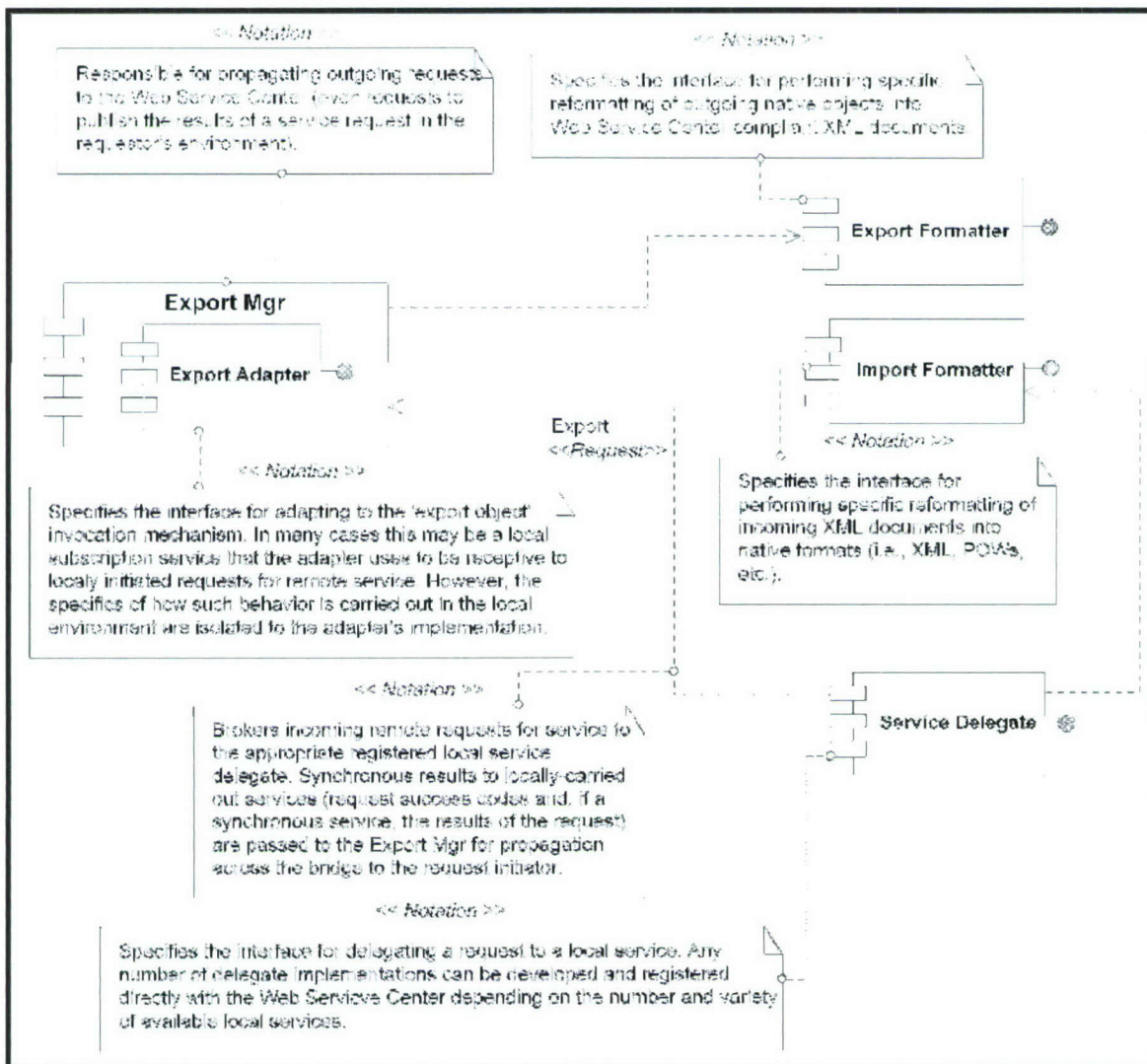


Figure 2 – Connector Framework Architecture

Outgoing communications are managed collectively by the Export Manager, Export Adapter, and Export Formatter. The main function of the Export Adapter is to employ the mechanism offered by the local client environment for becoming aware of, and obtaining outgoing communications (i.e., remote service requests or, in the case of a more bridge-isolated client, the local events that would trigger the connector to invoke a remote service). In either degree of awareness, a common client mechanism to facilitate this invocation is a client-level event service

capable of notifying interested parties (in this case, the Export Adapter) of the occurrence of certain local events. For more bridge-aware clients, such a mechanism may simply take the form of explicit method calls made to an extended Export Adapter implementation. However, the Connector Framework makes no demands on the level of client-intimacy of such a mechanism other than its existence.

Once the Export Adapter has received outgoing communications, it is passed to the Export Manager for generic *outgoing communications* processing. Such processing involves managing the necessary reformatting of the communication content into its XML equivalent. Recall that all direct interaction with the Service Center is XML-based. The details of this reformatting operation are completely encapsulated inside the client-specific implementation of the Export Formatter interface. While Service Center clients already capable of communicating in XML can avoid the overhead associated with this extra step, having such a reformatting capability allows non-XML capabilities to effectively utilize this interoperability solution in an architecturally organized manner. This again illustrates an underlying theme of this solution to limit constraints placed on system representation and format. Once the communication has been appropriately formatted into its XML equivalent, the Export Manager passes the content to the Service Center for brokering to the appropriate member of the Translator Pool.

Following this translation, the content is routed as incoming communication to the appropriate Service Delegate implementation of the target client's connector via the bridge's internal publish/subscribe facility. Adhering to the interface specified by the Connector Framework, each Service Delegate implementation essentially represents a *proxy*, or representative, for a local capability exposed to the remote world. It should be noted that by the time any incoming communication reaches a particular Service Delegate, the Service Center, via its Translation Pool, has already performed any necessary representational transformation ensuring that the target connector only receives content compliant with its client's native representation. Once a request is received by a Service Delegate, it is passed through the Import Formatter which converts the content to the target client's native format. Similar to the applicability of the Export Formatter, this step is only necessary if the native interface presented by the local client is not XML. It should also be noted that both the Export and Import Formatters do not perform the type of representational transformation undertaken by the translators housed within the Translator Pool. These formatters simply convert non-XML formats to their XML equivalents, and vice versa.

Once the request has been converted into the appropriate format, the Service Delegate invokes the native capability offered by its client to perform the requested service as well as manages the subsequent returning of any results to the Service Center as outgoing communications. Details of exactly how local capabilities are invoked and interacted with are fully encapsulated inside the client-specific implementation of the connector framework's Service Delegate interface. Such invocation may take a variety of forms including direct interface interaction or creation of a local event that in turn triggers the desired functionality. Regardless of the means of invocation, the functionality being requested may be at varying stages of formality. In other words, since the Service Delegates are essentially the representatives of an exposed capability, exactly what local functionality constitutes the particular service is encapsulated within the delegate itself and may

be in varying degrees of formalization. This is particularly useful when adapting legacy functionality to an interoperability paradigm.

The scenario presented above is, of course, most suitable where there is no native concept of interoperability outside the scope of a pre-determined set of systems. However, in the case where native capabilities are designed to operate in terms of an extendable open architecture, the role of Service Delegates can be reduced to managing the reformatting of communications in the case of non-XML systems, or, in the case where XML is supported, omitted completely. In the latter scenario the native capability would manage its own exposure to the Service Center but would still benefit from the virtual homogeneous representational environment promoted by the Interoperability Bridge facility.

#### 4. CONCLUSION

The Interoperability Bridge offers an effective means where by existing, perhaps loosely defined, system functionality can be adapted to operate within a dynamic interoperability paradigm. Through the use of Service Delegates, the details associated with directly interfacing with local system capabilities are encapsulated within client-specific code fragments and effectively isolated from reusable framework components. With flexibility as a fundamental theme, systems developed with such service-oriented concepts more integrated within their native design are able to avoid any undue overhead associated with such adaptation and exploit the functionality offered by the Interoperability Bridge in a more direct fashion.

The solution to interoperability among expressive environments presented in this discussion goes beyond traditional connective architectures by addressing the dramatic representational disparity typically exhibited by context-oriented systems. Rather than constraining interoperating environments to a common representation, albeit perhaps rich in expressiveness, the Interoperability Bridge provides a mechanism whereby potentially disparate representations can effectively be integrated to form a virtually homogeneous view of the world. As a result, interoperating systems can function within a cohesive service-oriented paradigm while still maintaining the natively precise perspectives critical to effective context-based decision-support.

#### REFERENCES

- [1] Cagle, K., M. Corning, J. Diamond, T. Duynstee, O. Gudmundsson, M. Mason, J. Pinnock, P. Spencer, J. Tang, A. Watt, J. Jirat, P. Tchistopolskii, and J. Tennison, "Professional XSL", Wrox Press Ltd., Birmingham, UK., 2001
- [2] Daconta M., L. Obrst and K. Smith, "The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management", Wiley, Indianapolis, IN., 2003
- [3] Ewalt D., "The Next Web", Information Week, October 2002, ([www.informationweek.com/story/IWK20021010S0016](http://www.informationweek.com/story/IWK20021010S0016))

- [4] Fikes R. and D. McGuinness, "An Axiomatic Semantics for RDF, RDF Schema and DAML+OIL", KSL Technical Report (KSL-01-01), October 2001, ([www.ksl.stanford.edu/people/dlm/daml-semantic/abstract-axiomatic-semantic.html](http://www.ksl.stanford.edu/people/dlm/daml-semantic/abstract-axiomatic-semantic.html))
- [5] Fowler M and K Scott, "UML Distilled: Applying the Standard Object Modeling Language", Addison-Wesley, Reading, Massachusetts, 1997.
- [6] Friedman-Hill, E., "JESS In Action", Manning Publications Co., Greenwich, CT, 2003
- [7] Garshol L. and G. Moore (eds.), "The XML Topic Maps (XTM) Syntax", JTC1/SC34:ISO 13250, July 22, 2002, ([www.y12.doe.gov/sgml/sc34/document/0328.htm](http://www.y12.doe.gov/sgml/sc34/document/0328.htm))
- [8] Giarratano J. and Riley G., "Expert Systems: Principles and Programming", 2<sup>nd</sup> Edition, PWS Publishing Company, Boston, MA.
- [9] Gil Y. and V. Ratnakar, "Markup Languages: Comparison and Examples", Information Sciences Institute, University of Southern California, TRELLIS project, 2002, ([www.isi.edu/expect/web/semanticweb/comparison.html](http://www.isi.edu/expect/web/semanticweb/comparison.html))
- [10] Graham S., S. Simeonov, T. Boubez, D. Davis, G. Daniels, Y. Nakamura, and R. Neyama, "Building Web Services with Java: Making Sense of XML, SOAP, WSDL, and UDDI", Sams Publishing, Indianapolis, IN, December 2001
- [11] Heflin J., R. Volz and J. Dale (eds.), "Requirements for a Web Ontology Language", W3C Working Draft, July 8, 2002, ([www.w3.org/TR/webont-req](http://www.w3.org/TR/webont-req))
- [12] Hendler J., T. Berners-Lee and E. Miller, "Integrating Applications on the Semantic Web", Journal of the Institute of Electrical Engineers of Japan, 122(10), October 2002, (pp.676-680).
- [13] Horrocks I., "DAML+OIL: A Description Language for the Semantic Web", IEEE Intelligent Systems, Trends and Controversies., 2002
- [14] Hunter D., C. Cagle, D. Gibbons, N. Ozu, J. Pinnock, and P. Spencer, "Beginning XML", Wrox Press Ltd., Birmingham, UK., 2000
- [15] Karsai G., "Design Tool Integration: An Exercise in Semantic Interoperability", Proceedings of the IEEE Engineering of Computer Based Systems, Edinburgh, UK, March, 2000
- [16] Manola F. and E. Miller (eds.), "RDF Primer", W3C Working Draft, March 19, 2002, ([www.w3.org/TR/2002/WD-rdf-primer-20020319/](http://www.w3.org/TR/2002/WD-rdf-primer-20020319/))
- [17] NASA, "CLIPS 6.0 Reference Manual", Software Technologies Branch, Lyndon B Space Center, Houston, Texas, 1992

[18] Pohl J., "Information-Centric Decision-Support Systems: A Blueprint for Interoperability", Office of Naval Research (ONR) Workshop hosted by the CAD Research Center in Quantico, VA, June 5-7, 2001

[19] Pohl J, A Chapman, K Pohl, J Primrose and A Wozniak, "Decision-Support Systems: Notions, Prototypes, and In-Use Applications", Technical Report, CADRU-11-97, CAD Research Center, Design Institute, College of Architecture and Environmental Design, Cal Poly, San Luis Obispo, CA, January, 1997

[20] Pohl K., "Perspective Filters As A Means For Interoperability Among Information-Centric Decision-Support Systems", Office of Naval Research (ONR) Workshop hosted by the CAD Research Center in Quantico, VA, June 5-7, 2001

[21] Simple Object Access Protocol (SOAP) Version 1.1; [www.w3.org/TR/soap](http://www.w3.org/TR/soap)

### AUTHOR BIOGRAPHY



*Kym J. Pohl is a senior software engineer and director with CDM Technologies Inc. in San Luis Obispo. His current focus is on agent-based, collaborative decision-support systems with particular interest in representation and collaborative architectures. Following an undergraduate degree in Computer Science he earned further Master degrees in Computer Science and Architecture. Over the past 20 years he has provided technical leadership in the design and development of a number of multi-agent decision-support systems for the US Department of Defense, including the Integrated Marine Multi-Agent Command and Control System (IMMACCS) for tactical command and control and the SEAWAY system for the coordination of logistical sea-based sustainment operations.*

# A TRANSLATION ENGINE IN SUPPORT OF CONTEXT-LEVEL INTEROPERABILITY

KYM J. POHL

*CDM Technologies Inc.*  
*San Luis Obispo, CA. 93401, USA*  
*805-541-3750 / Fax: 805-541-8296 / [kpohl@cdmtech.com](mailto:kpohl@cdmtech.com)*

**Abstract:** The need to support context-level interoperability is increasingly gaining importance in today's arena of semantic-oriented, decision-support systems. Unlike data-oriented exchange, such semantic interoperability must venture beyond the elementary communication of discrete data values and endeavor to translate between significantly more expressive, context-rich representations. Further, support of this level of interoperability must not require contamination of native perspective embedded within each participant's representation. The solution offered in this paper presents a service-oriented framework supporting an extensible set of translation paradigms to effectively connect expressive, ontology-based environments. Fundamental to this solution is the notion of a remote service request. Employing this metaphor as the basis for participant interaction allows each system within a universe of potentially diverse representations to interoperate as collections of invocatable services. Further, by transparently marshalling such requests between client and service representations, such translation engine offers each member of this multi-lingual reality the means to interoperate within the familiar confines of their native representations. Finally, the discussion concludes with an evaluation of this capability in terms of the Levels of Conceptual Interoperability Model (LCIM) for assessing degrees of interoperability.

**Keywords:** interoperability, decision-support, semantic mapping, LCIM

## 1. THE INTEROPERABILITY DILEMMA

The topic of semantic interoperability is increasingly gaining importance in today's information age. With the context-sensitive analytical demands placed on today's software-based decision-support systems [Pohl et al, 1997] it is more critical than ever for interaction among such systems to go beyond the simple sharing of data and venture into the realm of semantic exchange. Context-oriented systems, as opposed to data-centric systems, rely heavily on the contextual depth of the descriptions over which they operate. Contextual depth and semantic expression (i.e., property and relationship-rich descriptions of the entities and concepts representing a relevant set of domains) form the fundamental enabler for such systems to offer users meaningful assistance in the areas of resource-allocation, threat-analysis, and conflict-resolution to name a few. Whether interacting with other context-enabled systems or accepting data feeds from more data-centric systems, the need to communicate context as opposed to simply structure no doubt places a significantly higher burden on the representational depth of the overall exchange. Providing support for such context-centric interoperability is the topic of much research within academia and industry alike [Karsai, 2000].

Perhaps not surprisingly, it is the fundamental strength of context-centric, decision-support systems that poses the most challenging obstacle in this endeavor. This critical

enabler, and simultaneous nemesis, is *representational expressiveness*. As the term implies, the context-oriented approach to building decision-support systems endeavors to go beyond the classical *nuts and bolts* approach to representation (i.e., isolated chunks of typically numeric or string-based data with little or no inter-relationships and essentially void of any derivable meaning) and incorporate the potentially numerous relationships, constraints, and business rules that are needed for the more complex analysis inherent in agent-based, decision-support analysis. A critical aspect of such representational depth is *perspective*. The biases associated with how a particular entity or concept is *viewed* are exceedingly significant to the decision-making process. As such, perspective is a critical ingredient to truly effective context-oriented representation. When modeling, for example, a fundamental aspect of a transportation domain, a vehicle for instance, supporting the perception of this vehicle as a means for transporting cargo is very appropriate for a logistics support operations. However, for manufacturing operations, this same entity is perhaps more accurately represented as a sequenced collection of assembly stages. As such, in the scaled reality of today's multi-dimensional software-dependent decision-support environment, the eventual dilemma created by each context-based system essentially *seeing the world* according to its own personalized, albeit expressive, domain model (i.e., ontology) is the increased disparity that inevitably develops between the representations upon which these systems operate. In other words, the most empowering

ingredient in context-oriented computing also presents one of the most perplexing issues when such systems are required to interoperate [Pohl J., 2001; Pohl K., 2001].

To support meaningful collaboration within such decision-support environments there must be a translational component capable of mapping between expressive context in a manner that preserves the perspectives native to individual users, and does so in a transparent fashion. It is this focus that drove the design and development of the interoperability facility presented in this discussion. The following sections describe the criteria employed in developing a suitable solution to this paradigm, the various technologies exploited to realize its implementation, a discussion of how these pieces are combined to form an effective design capable of supporting expressive interoperability between context-oriented systems, and finally, an assessment of this interoperability platform in terms of the Levels of Conceptual Interoperability Model (LCIM) [Tolk and Muguira, 2003].

## 2. SOLUTION CRITERIA

To successfully address these issues, candidate technologies must satisfy several criteria. Each of these criteria is discussed in detail below.

One of the primary goals of any solution intended for repeated application to varying interoperability scenarios is the ability to be flexible. A key aspect to such adaptability is the clear separation of framework from application specifics. In other words, a candidate framework should support the various abstractions associated with translation-based interoperability among heterogeneous representations. Some of these abstracted concepts include the notion of translation itself, which should be broad enough to include an extendable variety of translation paradigms.

Further, this solution should identify and formalize the necessary interfaces to support adaptation of native client connectivity to the paradigm provided by the bridging framework. The latter requiring the necessary functionality and blueprint to effectively *adapt* client systems to the particular interoperability framework and interaction model presented by the solution.

Another significant property of any realistic solution intended for broad use is the support and promotion of industry standards. This is particularly significant when a high degree of reusability is intended. Accordingly, such a facility should center on industry-familiar technologies, standards, and tools. Not only does adherence to available standards offer a helpful guide in developing a particular solution but also aids in avoiding the unpleasant consequences associated with having to rework a problem

to fit an inflexible solution. This requires particular attention in a field where complexity and *one-off* solutions abound.

As stated in the previous section, a critical ingredient for effective interoperability among context-oriented systems is the preservation of native perspective. Successfully addressing this issue requires the ability to understand the subtleties inherent in such a concept (e.g., implied domain-specific constraints that have equally obscure and individual counterparts when considered from other domain-related perspectives). The considerations involved in translating between such perspectives often require a level of reasoning approaching the realm of expert systems [Giarratano and Riley, 2003]. In many respects, for the more complex context-oriented interaction a level of decision-support on par with solutions supporting multi-variable, complex problems is needed. In fact, in certain cases effective contextual translation may require the inclusion of the human decision maker for representing the higher-level logic not readily encodable within even today's cutting edge intelligent software.

And finally, an important, yet often overlooked, criterion for successfully addressing this interoperability problem is to not only focus on exchange of expressive context but to equally support less complex exchanges without incurring the overhead associated with support for the former. This is often the case where solutions targeting complex problems offer over-engineered and subsequently inefficient solutions for less taxing scenarios. The goal is to support a range of complexities and to limit any incurred overhead to situations where it cannot be avoided.

## 3. SUITABLE TECHNOLOGIES AND DESIGN PATTERNS

To successfully address the challenging criteria set forth above several enabling technologies and approaches were investigated. First among these is a Service-Oriented Architecture (SOA) for operation in both standalone and web-enabled forms [Friedman-Hill, 2003; Ewalt, 2002; Graham et al, 2001; Heflin et al; 2002; Hendler et al, 2002; Horrocks, 2002]. In this manner, services can be registered, consequently discovered, and eventually invoked via standard registry lookup and interaction protocols. Further, by defining the functional topology in terms of invocable services, the classical notion of system boundaries is replaced with the more dynamic notion of composable services effectively expanding and contracting system boundaries as the need arises.

The eXtensible Markup Language (XML) [Hendler et al, 2002; Hunter et al, 2000] together with its Extensible Stylesheet Language Transform (XSLT) [Hunter et al,

2000] language counterpart are two additional technologies applicable to the domain of translation. XML provides a flexible means of defining structure through the use of syntactical tags. XML schemas can be developed to describe the structural aspects of entities, notions, and concepts. Receivers can process XML statements based on these schemas in an interpretational manner. The result, although accompanied by a significant caveat, is a means whereby software components can process incoming content based on a previously unknown representation. However, it should be noted that such discovery is limited to structural characteristics and does not include the discovery of semantics, *or meaning*, vital in context-oriented exchange. Even considering a meta-level schema describing the domains of concepts, rules, and implications, there is still the requirement for processors of such content to be equipped with a pre-conceived understanding of the semantics of even these abstracted domains. At some point, the semantics need to be adequately represented beyond simple structure. That said, however, structural detection does play a significant role in the eventual goal of true contextual discovery but should be viewed with a strong distinction between structure and semantics.

Perhaps one of the more significant additions to the world of tag-based schema languages is the ability to formally describe inter-schema transformation rules in an XML-friendly syntax. XSLT is a language that can be used to describe exactly how content based on one XML schema can be mapped into another. Whether defined statically or more dynamically at runtime, XSL transforms can be effectively applied for straightforward, property-to-property translations.

Although translation at this level is useful, for the more complex transformation inherent in context-oriented representations a more powerful paradigm is required. More extensive reasoning capabilities can be realized through the use of inference engine-based technology. Similar to XSLT, inference engine-based translation represents transformation logic as sets of managed rules. However, as opposed to the aforementioned XSLT mappings, these rule sets can embody significantly more

complex logic supporting extensive condition-based pattern matching and context-based inferencing. Some examples of rule-based inference engines include the CLIPS expert system shell developed by NASA [NASA, 1992] and a somewhat similar, although more easily embeddable within Java-based environments, JESS inference engine developed by Sandia Laboratories [Friedman-Hill, 2003]. In either case, complex transformation logic can be implemented as pseudo expert systems applying various levels of reasoning to determine the appropriate transformation(s) to apply given an expressive contextual fragment. This added level of sophistication is particularly useful where translated subject matter is bound by different constraints as it moves between representations. Under these circumstances, enforcement of such constraints may require a level of decision-support capable of reorganizing the particular subject matter in a manner that produces the appropriate state upon insertion into the target representation.

#### **4. A SOLUTION: THE INTEROPERABILITY BRIDGE**

The solution presented in this discussion takes the form of a configurable, service-oriented interoperability framework capable of supporting an extensible set of translation paradigms. Further, the Interoperability Bridge, or *Bridge* for short, introduces an overarching interaction model founded on a *remote service request* metaphor. The orientation toward this type of exchange model allows each interoperating system to view its counterparts as collections of composable services, invocable using language (i.e., representation) native to the caller. The resulting interoperability environment promotes a decoupled architecture requiring no pre-conceived notion of participant identity other than the conceptual services a client chooses to expose.

The main Bridge architecture is essentially comprised of three components. These include the Service Center, an extensible Translator Pool, and individualized Client Connectors. The following section discusses each of these components in more detail.

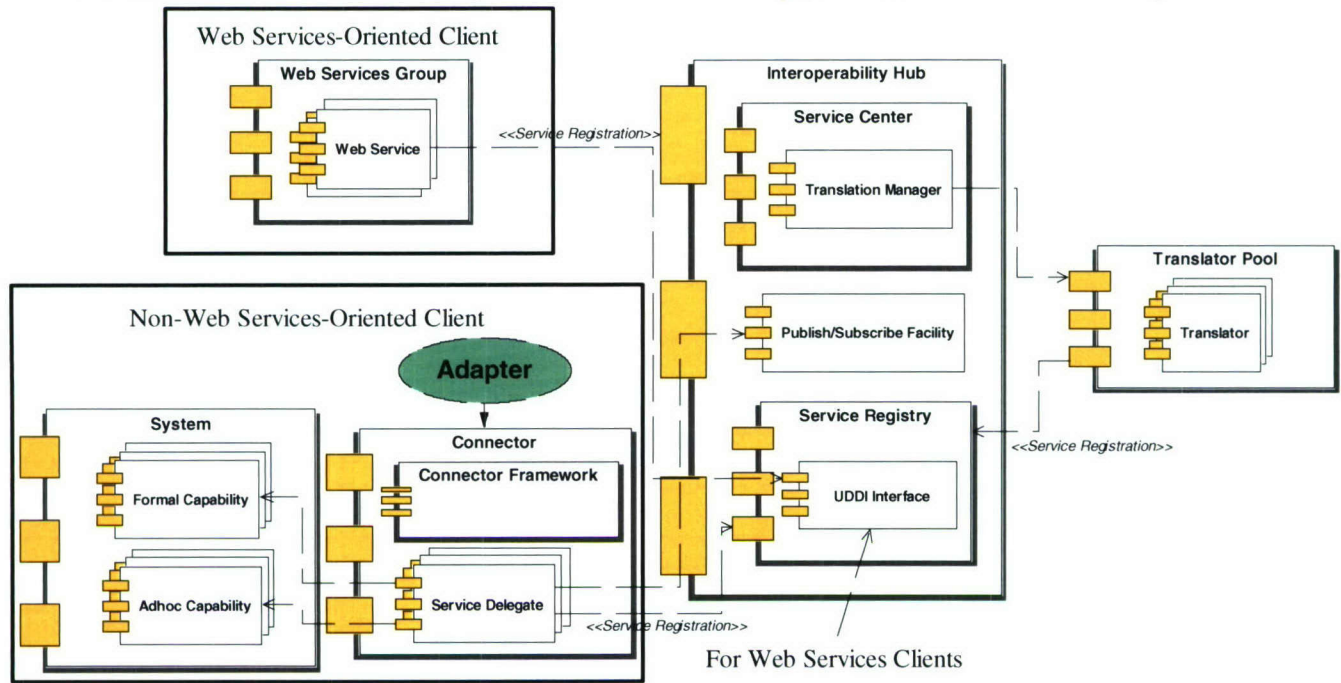


Figure 1 – The Interoperability Bridge Component Architecture

### The Service Center

The Service Center provides the overall management required to orchestrate the reception and processing of client requests. As the main interface for Bridge clients, this component is exposed in both a standalone and web service manner. When deployed as a web service, the Bridge adheres to the standard web service design pattern [Graham et al, 2001]. In its standalone form, the Service Center is exposed as a programmatic interface. Figure 1 provides a Unified Modeling Language (UML) [Fowler and Scott, 1997] illustration of how the Service Center, in this case configured as a web service, promotes a service-oriented architecture allowing clients to expose both formalized and ad hoc capabilities as composable services available for invocation.

Apart from providing the main client interface of the Bridge, the Service Center also acts as a request broker effectively routing inter-client communication (i.e., service requests and subsequent responses) to the appropriate member of the Translator Pool for effective inter-dialect mapping (i.e., mapping between client representations). The Service Center accomplishes this task in the form of a Translation Manager. The Translation Manager embedded within the Service Center is responsible for discovering and engaging suitable translation services (i.e., translators) to perform the required representational mappings. To further support the web service interaction model promoted at the main client interface level, the Translation Manager is capable of interacting with registered translators in either a standalone or web service form.

### The Translator Pool

Collectively known as the Translator Pool, and the second of the three primary components comprising the Bridge architecture, the particular set of translators registered with the Bridge at any point in time determines the available translation paradigms available to interoperating clients. Further, the breadth and depth of interoperability available within these paradigms is determined by the various translation rule sets these translators are equipped with. As with membership within the Translator Pool itself, such configuration occurs in a dynamic fashion as translation needs change and desirable rule sets become available. As a translator's scope of language mappings change (i.e., a translator being quipped with additional or modified mapping rules), such translation capability is communicated to the Translation Manager in terms of both the source and target XML Schema Definition (XSD) that the associated rule set can map between.

### The Client Connector

The various Client Connectors that effectively *connect* each client to the Bridge complete the primary Bridge architecture. Each client intending to invoke the services native to the Bridge (i.e., publication, query, subscription) or the remote services offered by other clients does so via its specific Client Connector. Based on a reusable framework provided as part of the Bridge development and execution package, a Connector is built for each type of candidate bridge client in accordance with that particular connectivity model. In essence, these connectors effectively *adapt* a client to seamlessly

operate in terms of the interaction model promoted by the Bridge. This approach has a very significant advantage over more intrusive connectivity models. Essentially acting as a *proxy* for its specific client, the Connector effectively isolates functionality native to the client from any knowledge of the Bridge itself. During invocation of a client's service, the receiving client's Connector processes the request into a series of native function calls. For outgoing requests, clients have a choice as to their level of awareness of the Bridge. For client's wishing to initiate remote requests, such invocation logic may be woven into the functional fabric of the particular client, or, may be solely isolated within the implementation of the Connector thus preserving a client's decoupled nature. The latter of these two configuration extremes exploits the *delegate* design pattern implemented within the Connector architecture. For more isolated clients such delegates can react to locally occurring events, issuing remote service requests on an as needed basis. In either case, the results of inter-client service invocations can be asynchronously returned as service requests directed toward the issuer's particular publication service, in whatever form that capability may be.

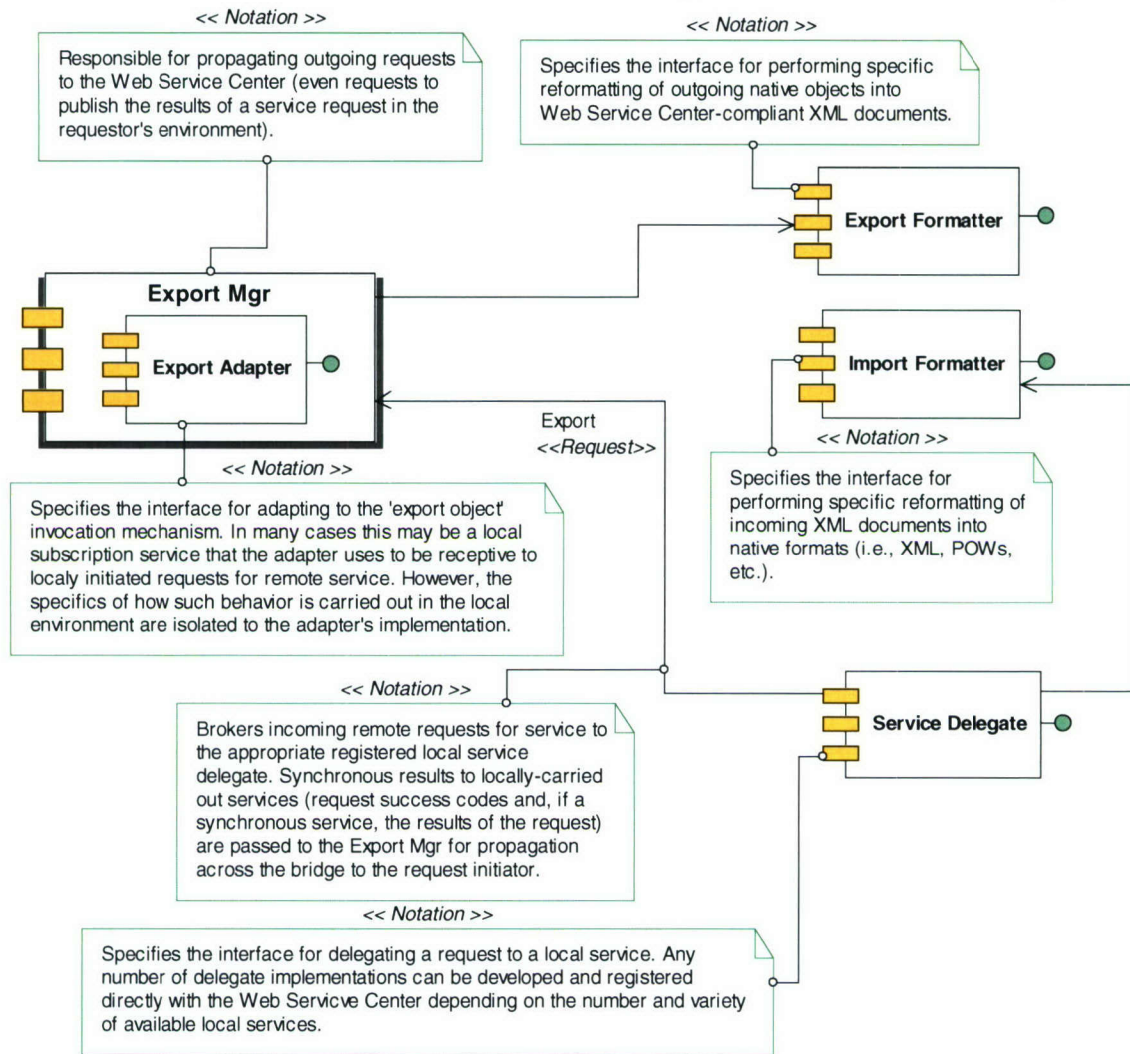
### The Client Connector Framework

To assist in coping with differing flavors of client connectivity, the Interoperability Bridge platform offers a framework for building and managing Client Connectors. This framework consists of well-defined interfaces, functional modules, and a formalized design pattern for integrating these components with client-specific behavior. Figure 2 illustrates the architecture of the Connector Framework along with the various interfaces on which client-specific concerns are implemented. To provide a better understanding of how these Connectors operate in conjunction with both client and Bridge functionality, the following section provides a brief illustration of the Connector workflow involved in processing a typical client service request.

Referencing the Connector architecture presented in Figure 2, outgoing communications are managed collectively by the Export Manager, Export Adapter, and Export Formatter. The main function of the Export Adapter is to employ the mechanism offered by the local client for being notified of, and subsequently receiving, outgoing communications (i.e., remote service requests or, in the case of a less Bridge-aware client, the local events that would indirectly trigger the Connector to invoke a remote service). In the case of a more isolated

client, a common mechanism to facilitate indirect invocation may be some sort of local event service capable of notifying interested parties, in this case the Export Adapter, of certain events. For more Bridge-aware clients, such a mechanism may take the more direct form of explicit method calls issued to an extended Export Adapter implementation.

Once the Export Adapter has received outgoing content, it is passed to the Export Manager for processing. Such processing involves managing the necessary reformatting of the communication content into its XML equivalent. Recall that all direct interaction with the Service Center is XML-based. The details of this reformatting operation are completely encapsulated inside the client-specific implementation of the Export Formatter interface. While Service Center clients already capable of communicating in XML can avoid the overhead associated with this extra step, having such a reformatting capability allows non-XML capabilities to effectively utilize the Bridge in an architecturally organized manner. This again illustrates an underlying theme of this solution to limit constraints placed on system representation and format. Once the communication has been appropriately formatted into its XML equivalent, the Export Manager passes the content to the Service Center for brokering to the appropriate member of the Translator Pool. Following this translation, the content is routed as incoming subject matter to the appropriate Service Delegate implementation of the target client's Connector via the bridge's internal publish/subscribe facility. Adhering to the interface specified by the Connector Framework, each Service Delegate implementation essentially represents a *proxy*, or representative, for a local capability exposed to the remote world. It should be noted that by the time any incoming content reaches a particular Service Delegate, the Service Center, via its Translation Manager, has already performed any necessary representational transformation ensuring that the target connector only receives content compliant with its client's native representation. Once a Service Delegate receives a request, it is passed through the Import Formatter that converts the XML-based content to the target client's native format. Similar to the applicability of the Export Formatter, this step is only necessary if the native interface presented by the receiving client is not XML. It should also be noted that both the Export and Import Formatters do not perform the type of representational transformation undertaken by the translators housed within the Translator Pool.



**Figure 2 – Connector Framework Architecture**

Once the request has been converted into the appropriate format, the Service Delegate invokes the native capability offered by its client to perform the requested service, as well as manages the subsequent returning of any results to the Service Center as outgoing content. Details of exactly how local capabilities are invoked and interacted with are fully encapsulated inside the client-specific implementation of the connector framework's Service Delegate interface. Such invocation may take a variety of forms including direct interaction or creation of a local event that indirectly triggers the desired client functionality. Regardless of the means of invocation, the functionality being requested may be at varying stages of formality. In other words, since the Service Delegates essentially function as the representatives of an exposed capability, exactly what local functionality constitutes the particular service is encapsulated within the delegate itself and may be in varying degrees of formalization. This is particularly useful when adapting loosely organized functionality to a more formalized, service-oriented interoperability paradigm.

The scenario presented above is most suitable where there is no native concept of interoperability outside the scope of a pre-determined set of systems. However, in the case where native capabilities are designed to operate in terms of an extensible, open architecture, the role of the Service Delegates can be reduced to managing the reformatting of communications in the case of non-XML systems, or, in the case where XML is supported, omitted completely. In the latter scenario the native capability would manage its own exposure to the Service Center but may still benefit from the virtual homogeneous environment promoted by the Interoperability Bridge facility (i.e., each Bridge client is provided the impression that their native language is the interoperability standard when in fact, each client is communicating in terms of their specific dialect).

## 5. THE INTEROPERABILITY BRIDGE AND SUPPORTABLE LEVELS OF CONCEPTUAL INTEROPERABILITY (LCIM)

LCIM offers a multi-tiered set of criteria for assessing the degree of interoperability achieved between software systems. Evolving from efforts in the world of modeling and simulation, LCIM presents the following 7-tier interoperability assessment model [Tolk and Muguira, 2003]:

- **Level 0:** Stand-alone systems having essentially no interoperability.
- **Level 1:** Achieving *technical interoperability*, a communication protocol exists for exchanging data between participating systems. Further, a communication infrastructure is established allowing exchange of bits and bytes with underlying network and communication protocols unambiguously defined.
- **Level 2:** Achieving *syntactic interoperability*, there exists a common structure to exchange information, i.e., a common data format is applied. At this level, a common protocol to structure the data is used with the format of the information exchange unambiguously defined.
- **Level 3:** Applying a common information exchange reference model, a level of *semantic interoperability* is reached. At this level, the meaning of the data is shared and the content of the information exchange requests are unambiguously defined.
- **Level 4:** *Pragmatic interoperability* is achieved when the interoperating systems are aware of the methods and procedures that each other are employing. In other words, the use of the data, or the context of its application, is effectively *understood* by the participating systems. Further, the context in which the information is exchanged is also unambiguously defined.
- **Level 5:** A level of *dynamic interoperability* is achieved when interoperating systems are able to comprehend the various changes in state (i.e., assumptions and constraints) made by each other as execution progresses. Achieving this level of interoperability allows participating systems to effectively take advantage of such knowledge to their benefit.
- **Level 6:** *Conceptual interoperability*, the highest level of LCIM interoperability, is achieved when the conceptual model (i.e. the assumptions and constraints of the purposeful abstraction of reality) become semantically aligned. Further, this level of interoperability requires that conceptual models be

documented in their implementation-independent form and subsequently made available to engineers.

As opposed to an actual instance of a particular representational model, the Interoperability Bridge discussed in this paper offers a platform where such models can be defined and executed. As such, LCIM assessment is performed in terms of the various paradigms the Bridge is design to support.

Whether deployed in its standalone or web service form, the Interoperability Bridge provides an underlying communication infrastructure that, through customization of the adaptive Connectors, effectively connects both homogeneous and heterogeneous clients. This capability provides the *technical interoperability* specified in LCIM Level 1.

Both *syntactic* and *semantic interoperability*, LCIM Levels 2 and 3 respectively, is achieved once the Interoperability Bridge is equipped with the appropriate mapping models that effectively tie together client representations. As described earlier, the Bridge provides an expandable set of translation paradigms suitable for both simple and complex inter-representational mapping. The Bridge also provides a model for both constructing and integrating addition translation engines allowing for incorporation of emerging mapping technologies as they arise.

LCIM Level 4, *pragmatic interoperability*, requires exposure of client capabilities to other interoperating clients. The Interoperability Bridge supports this task by defining client interaction in terms of a *remote service request* metaphor. In this manner, clients are able to not only expose available functionality to interested clients but can present such capabilities as *conceptual* services that in actuality may be transparently comprised of various collections of formalized or more ad hoc native functionality.

Considering the degree of expressive interoperability manageable by the Bridge, cross-client awareness and comprehension of each other's changes in state (i.e., LCIM Level 5) is supportable but essentially a matter of the degree to which such aspects are represented in the particular interoperability model the Bridge is configured to manage. In other words, while the Bridge has no native specification for such inter-client awareness it, nonetheless, does present a suitable environment to support an interoperability model equipped to represent and convey such notions across its clientele. As mentioned at the beginning of this section, as opposed to a specific instance of an interoperability model, the Bridge offers a platform enabling varying degrees of inter-client exchange constrained only by the scope and sophistication of the configured interoperability model and the ability of its clientele to process such representation. Further, in support of the higher levels of client awareness and comprehension associated with

LCIM level 5, the Bridge provides a service-oriented client interaction model together with an extendable set of translation paradigms suitable for managing notions as sophisticated as domain-specific assumptions and constraints.

Although the Bridge is currently only concerned with the interoperability model in its implementation form, it would be interesting to consider the potential for supporting such a model in the more conceptual form specified in LCIM Level 6. Although the Bridge can currently process models specified in terms of the Unified Modeling Language (UML) methodology, the Bridge makes no distinction as to the conceptual nature of such descriptions. In other words, such models are processed by the Bridge as the literal language for inter-client exchange and not the more abstracted form on which various implementations of such a model could be based. To not only acknowledge such a distinction, but to also provide a meaningful level of support, an ability to relate implementation model fragments to their conceptual model counterparts would appear to be helpful. Once such a relationship was understood, there may be potential for the Bridge to exploit this knowledge to automatically determine relationships existing between various conceptual model implementations. In this manner, there would be no need to develop, and consequentially configure the Bridge, with explicit mappings between implementation models. Such connections would be developed and dynamically maintained by the Bridge at run time. Taking this capability to the next level, if the Bridge could, in fact, comprehend the conceptual nature of an interoperability model, perhaps in terms of a native *concept* ontology, it would appear plausible that such awareness could begin to allow the Bridge to elevate its analysis of client models to their conceptual form. Such capability would set the stage for focusing the Bridge's support for dynamically determining cross-client interoperability to the conceptual level drawing concept-based relationships between native client models in their conceptual form.

## 6. CONCLUSION

The Interoperability Bridge offers an effective means where by existing, perhaps loosely defined, system functionality can be adapted to operate within a dynamic interoperability paradigm. Through the use of Service Delegates, the details associated with directly interfacing with local system capabilities are encapsulated within client-specific code fragments and effectively isolated from reusable framework components. With flexibility as a fundamental theme, systems developed with such service-oriented concepts more integrated within their native design are able to avoid any undue overhead associated with such adaptation and more readily exploit the functionality offered by the Interoperability Bridge.

The solution to interoperability among expressive environments presented in this discussion goes beyond traditional connective architectures by addressing the dramatic representational disparity typically exhibited by context-oriented systems. Rather than constraining interoperating environments to a common representation, albeit perhaps rich in expressiveness, the Interoperability Bridge provides a platform whereby potentially disparate representations can effectively be integrated to form a virtually homogeneous view of the world. As a result, interoperating systems can function within a cohesive service-oriented paradigm while still maintaining the natively perspectives critical to effective context-based decision-support.

## REFERENCES

- Cagle, K., M. Corning, J. Diamond, T. Duynstee, O. Gudmundsson, M. Mason, J. Pinnock, P. Spencer, J. Tang, A. Watt, J. Jirat, P. Tchistopolskii, and J. Tennison, "Professional XSL", Wrox Press Ltd., Birmingham, UK., 2001
- Daconta M., L. Obrst and K. Smith, "The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management", Wiley, Indianapolis, IN., 2003
- Ewalt D., "The Next Web", Information Week, October 2002,  
([www.informationweek.com/story/IWK20021010S0016](http://www.informationweek.com/story/IWK20021010S0016))
- Fowler M and K Scott, "UML Distilled: Applying the Standard Object Modeling Language", Addison-Wesley, Reading, Massachusetts, 1997.
- Friedman-Hill, E., "JESS In Action: rule-based systems in java", Manning Publications Co., Greenwich, CT, 2003
- Giarratano J. and Riley G., "Expert Systems: Principles and Programming", 2<sup>nd</sup> Edition, PWS Publishing Company, Boston, MA.
- Gil Y. and V. Ratnakar, "Markup Languages: Comparison and Examples", Information Sciences Institute, University of Southern California, TRELIS project, 2002,  
([www.isi.edu/expect/web/semanticweb/comparison.html](http://www.isi.edu/expect/web/semanticweb/comparison.html))
- Graham S., S. Simeonov, T. Boubez, D. Davis, G. Daniels, Y. Nakamura, and R. Neyama, "Building Web Services with Java: Making Sense of XML, SOAP, WSDL, and UDDI", Sams Publishing, Indianapolis, IN, December 2001

Heflin J., R. Volz and J. Dale (eds.), "Requirements for a Web Ontology Language", W3C Working Draft, July 8, 2002, ([www.w3.org/TR/webont-req](http://www.w3.org/TR/webont-req))

Hendler J., T. Berners-Lee and E. Miller, "Integrating Applications on the Semantic Web", Journal of the Institute of Electrical Engineers of Japan, 122(10), October 2002, (pp.676-680).

Horrocks I., "DAML+OIL: A Description Language for the Semantic Web", IEEE Intelligent Systems, Trends and Controversies., 2002

Hunter D., C. Cagle, D. Gibbons, N. Ozu, J. Pinnock, and P. Spencer, "Beginning XML", Wrox Press Ltd., Birmingham, UK., 2000

Karsai G., "Design Tool Integration: An Exercise in Semantic Interoperability", Proceedings of the IEEE Engineering of Computer Based Systems, Edinburgh, UK, March, 2000

NASA, "CLIPS 6.0 Reference Manual", Software Technologies Branch, Lyndon B Space Center, Houston, Texas, 1992

Pohl J., "Information-Centric Decision-Support Systems: A Blueprint for Interoperability", Office of Naval Research (ONR) Workshop hosted by the CAD Research Center in Quantico, VA, June 5-7, 2001

Pohl J, A Chapman, K Pohl, J Primrose and A Wozniak, "Decision-Support Systems: Notions, Prototypes, and In-Use Applications", Technical Report, CADRU-11-97, CAD Research Center, Design Institute, College of Architecture and Environmental Design, Cal Poly, San Luis Obispo, CA, January, 1997

Pohl K., "Perspective Filters As A Means For Interoperability Among Information-Centric Decision-Support Systems", Office of Naval Research (ONR) Workshop hosted by the CAD Research Center in Quantico, VA, June 5-7, 2001

Tolk A. and Muguira J. A., "The Levels of Conceptual Interoperability Model (LCIM)." Proceedings IEEE Fall Simulation Interoperability Work-shop, IEEE CS Press , 2003

*systems with particular interest in expressive representation and collaborative architectures. Following an undergraduate degree in Computer Science he earned Master's degrees in both Computer Science and Architecture. Over the past 20 years he has provided technical leadership in the design and development of a number of context-oriented, decision-support systems for the US Department of Defense, including the Integrated Marine Multi-Agent Command and Control System (IMMACCS) for tactical command and control, the Joint Force Collaborative Toolkit (JFCT) supporting logistics planning for seabasing operations, and the TRANSWAY decision-support system providing theater-to-theater logistics planning and visibility.*

#### AUTHOR BIOGRAPHY



*Kym J. Pohl is a senior software engineer and director of CDM Technologies, Inc. in San Luis Obispo, California. His current focus is agent-based, collaborative decision-support*

## IV.C.2: Results of Research Projects

## **Molecular sensors and defenses against ultraviolet radiation**

Project Investigator:

Nikki L. Adams  
Biological Sciences  
California Polytechnic State University  
San Luis Obispo, California

## A. ABSTRACT

Most organisms are exposed to damaging levels of ultraviolet radiation (UVR, 290-400 nm). The overall goals of this study were to identify molecular targets of UVR-induced damage in cells and to quantify natural sunscreens (mycosporine-like amino acids, MAAs) in local marine organisms. Two Specific Aims were addressed by our research. **Aim 1:** Identify molecular targets (markers) of radiation-induced damage that lead to perturbations in cell division using sea urchin embryos as a model system. We continued work examining effects of UVR on a potential marker for radiation-induced damage in cells, checkpoint kinase 1 (Chk1), in local purple sea urchins, *Strongylocentrotus purpuratus* (SpChk1). We utilized early *S. purpuratus* genomic DNA predictions for Chk1 to design primers, clone and sequence SpChk1. SpChk1 is >60% homologous with mammalian Chk1s and is more similar to mouse Chk1 than any other organism tested. We used the resulting sequence to design a synthetic peptide and create antibodies against SpChk1 to help us determine the function of Chk1 in sensing radiation-induced damage sea urchin embryos. Both antibodies recognize a protein, which appears to be Chk1 in *S. purpuratus* embryos. In addition, this project facilitated our involvement in the Sea Urchin Genome Consortium and co-authorship by Adams in the ultimate publication of the map of the sea urchin genome (The Sea Urchin Genome Consortium, 2006). This map of the *S. purpuratus* genome provides a valuable resource for identifying molecular targets or regulators of environmental stress and demonstrates that the majority of sea urchin genes are more closely related to humans than they are to invertebrates. In most cases, sea urchins contain a single copy of genes and proteins that are redundant in humans and thus, provide a simple system for studying and understanding basic cellular processes and biomedical research. **Aim 2:** Identify and quantify MAAs in local sea urchins, *S. purpuratus*, to create a foundation for ultimately examining whether the targets identified in aim 1 are protected by MAAs and whether MAAs can be developed as broadband sunscreens. We demonstrated that *S. purpuratus* contain a suite of natural sunscreens (MAAs) in their ovaries and testes, but that the concentrations of MAAs vary within and among collection sites. We plan to continue both tracks of research to benefit marine science and the biotechnology industry by identifying molecular markers of radiation-induced stress and provide a means for testing the effectiveness of natural and commercial sunscreens.

## B. BACKGROUND AND GOALS OF RESEARCH

Depletion of stratospheric ozone has increased the amount of solar ultraviolet radiation (UVR, 290-400 nm, specifically UVB, 290-320 nm) reaching the Earth's surface, making it increasingly important to understand effects of UVR on organisms (Madronich et al., 1998). Solar radiation is the primary driving force for most processes in the ocean (Whitehead and De Mora, 2000). Nevertheless, UVR harms aquatic organisms such as embryonic stages of marine invertebrates that float at the ocean's surface or live in shallow water (Adams and Shick, 1996; 2001; Gleason and Wellington, 1996; Häder et al., 1998; Karentz and Bosch, 2001). Damaging levels of solar UVR can penetrate to at least 20 m in highly transparent water, 20-30 m in Antarctic waters (Smith et al. 1992), and to several meters in organically rich coastal seawater (Smith and Baker, 1979; Booth and Morrow; 1997; Adams et al., 2001). Assessment of UVR's impact on life histories at the ecological and cellular levels is particularly important because factors influencing early development will determine adult population characteristics.

UVR causes direct damage to DNA, proteins, and membrane lipids and indirect damage by forming reactive oxygen species (ROS) (Tyrrell, 1991). UV-irradiation of cells causes dosage-dependent delays in cell division and developmental abnormalities. These abnormalities may be

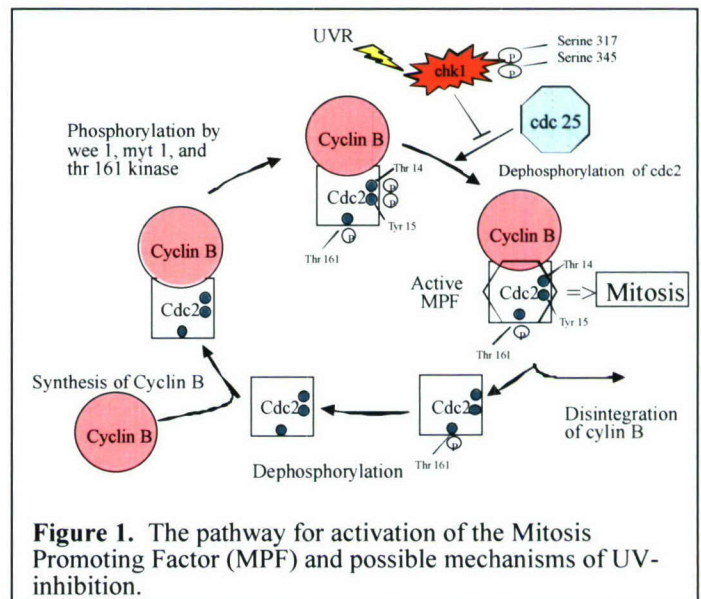
viewed microscopically, but it is important to understand their cellular and molecular origin by identifying the molecular targets of UVR.

One of the earliest model systems for study of this phenomenon at the cellular level was sea urchin eggs and embryos (Adams and Shick 1996; 2001; Rustad, 1971). In addition, among the non-chordate deuterostomes, the sea urchins are the primary target for research in genomics because they are the only non-chordate deuterostomes that serve as a major research model (Davidson et al., 2004). At least two highly conserved and critically important proteins have been discovered in sea urchins and revolutionized the fields of cell cycle and cytoskeletal research (cyclins and dynein). This year, the sea urchin genome was sequenced and annotated (The Sea Urchin Genome Consortium, 2006). Genomic data demonstrate that sea urchins share many genes with humans and that the majority of sea urchins genes are more closely related to mammals than they are to other invertebrates. In most cases, sea urchins contain a single copy of genes and proteins that are redundant in humans and thus, provide a simple system for studying basic cellular processes, including biomedical applications.

In all eukaryotes, the sequential formation, activation, and deactivation of a series of structurally regulated serine/threonine protein kinases regulate progression of cells through the cell cycle (Morgan, 1997). These enzymes consist of a regulatory subunit, a cyclin, and a catalytic subunit, Cdc2, a cyclin dependent kinase (cdk), which combine to form Mitosis-Promoting Factors (MPFs). Post-translational modification regulate activity of MPFs (Fig.1). This complex is the key regulator of cell cycle checkpoints known as the G2/M (mitosis) transition (Murray and Kirschner, 1991).

Upstream kinases (add phosphates) and phosphatases (remove phosphates) regulate MPF activity (Fig. 1). Upon activation of sea urchin eggs, protein synthesis begins and Cyclin B is synthesized. Cdc2 is present in eggs and binds to the newly synthesized Cyclin, to form an inactive MPF. Once the MPF is formed, two proteins, Myt 1 and Wee1, phosphorylate Cdc2 at two inhibitory sites, Tyr 15 and Thr 14. Then, Thr 161 kinase, phosphorylates Cdc2 at an activating site, Thr 161 (Coleman and Dunphy, 1994; Suprywicz et al., 2000). Accumulation of the repressed form of MPF to a threshold level creates precise control over entry into mitosis. Cdc25 (a phosphatase) is responsible for removing the inhibitory phosphates at Thr 14 and Tyr 15 activating MPF and propelling the cell into mitosis (Peng et al., 1997). If any of these events are delayed, mitosis is stalled (Fig. 1). This pathway is highly conserved among organisms and cell types, including echinoderm embryos (Meijer et al., 1989; Okano-Uchida et al., 1998; Suprywicz et al., 2000).

Dephosphorylation of Cdc2 at Tyr 15 and of Cdc25 at Ser 216 or S287 is inhibited by UVR in human keratinocytes (Gabrielli et al., 1997; Herzinger et al., 1995) or general damage to DNA (Furnari et al., 1997; Sanchez et al., 1997). Cdc25 is inhibited via phosphorylation at Ser 216 or S287 by checkpoint kinase 1 (Chk1). In mammalian cells, UV-irradiation and genotoxic stress activate Chk1, by phosphorylation of Ser 317 and 345 (Zhao et al., 2001). Examination of Chk1 phosphorylation may provide the most direct marker of UV-induced mitotic delays documenting molecular damage by UVR. The tight conservation of this pathway between sea urchins and mammalian cells indicate that this is likely to be the case in sea urchin embryos.



**Figure 1.** The pathway for activation of the Mitosis Promoting Factor (MPF) and possible mechanisms of UV-inhibition.

Our preliminary evidence support this hypothesis and indicate that further investigation of effects of UVR on this pathway in sea urchin embryos may allow us to more fully understand the mechanism of UV-induced cleavage delays.

Our previous experiments indicate dephosphorylation of Cdc2 at Tyr 15 is delayed in sea urchin zygotes that are exposed to artificial UVR. Once mitosis has progressed, Cdc2 becomes dephosphorylated at Thr 161 and is ready for another round of cell division. Thr 161 is phosphorylated similarly in embryos of both treatments, but that its dephosphorylation is delayed in UV-irradiated embryos, what we would expect if cell mitosis had been delayed. The delay in dephosphorylation of Cdc2 correlates precisely with the delays in cell division in these batches of embryos (Adams, in preparation). We hypothesize that activity (phosphorylation) upstream proteins in this pathway, including Cdc25 and Chk1 activity will be acted by UVR in sea urchin embryos.

Shallow-dwelling organisms have natural defenses and behaviors that may reduce exposure to solar UVR, particularly during periods of peak irradiance (Adams, 2001; Pennington & Emlet, 1986). Phototrophic organisms that rely directly on sunlight, or planktonic and sessile organisms that cannot avoid solar UV-exposure remain exposed to potentially damaging levels of solar UVR (Jerlov, 1950; Smith & Baker, 1979). Fortunately, they may receive protection by accumulating mycosporine-like amino acids (MAAs) that absorb UVR and may dissipate its energy harmlessly (Dunlap & Shick, 1998; Dunlap & Chalker, 1986; Shick et al., 2000).

MAAs are a suite of water-soluble compounds with absorption maxima ranging from 309 to 360 nm. They have overlapping absorption ranges that extend across much of the spectrum of UVR, creating effective broad-spectral sunscreens (Dunlap & Shick, 1998). They are ubiquitous among marine organisms and they protect sea urchin embryos against UV-induced delays in cell division and developmental abnormalities (Adams and Shick, 1996; 2001). It is still unclear, however, which molecular targets are affected by UVR in sea urchin embryos leading to cleavage delay and developmental abnormalities and which targets are protected by MAAs.

MAA concentration varies among tissues, often being highest in the tissues most exposed to UVR (Chioccare et al., 1986; Dunlap et al., 1989; Shick et al., 1992; Adams & Shick, 1996; 2001). MAA concentrations in corals and macroalgae are positively correlated with natural levels of UVR (Dunlap & Shick, 1998; Franklin et al., 1999; Karsten et al., 1998), so that organisms may regulate concentrations of UV-absorbing compounds in proportion to the amount of UVR they experience in their environment (Dunlap et al., 1986; Gleason, 1993; Shick et al., 1996; 1999). Red macroalgae have been studied extensively and shown to specifically regulate their concentrations of MAAs in relationship to UV-exposure (Franklin et al., 1999; Karsten et al., 1998). Thus, they have the potential to serve as bio-sensors of UV-exposure. Red algae studied to date contain a wide-array of MAAs and have also been used in experiments demonstrating dietary accumulation of MAAs from algae in marine invertebrates (Carroll and Shick, 2000; Adams and Shick 1996; 2001).

The synthesis of MAAs is restricted to bacteria, cyanobacteria, fungi, and algae (Favre-Bonvin et al., 1987; Shick et al., 1999). Marine consumers such as sea urchins and nudibranchs are unable to synthesize the compounds themselves and acquire MAAs from their diets of algae (Adams & Shick, 1996; Carefoot et al., 1998; Carroll & Shick, 1996). Adult sea urchins, *Strongylocentrotus droebachiensis*, sequester MAAs in their eggs, but not sperm (Adams & Shick, 1996). *S. droebachiensis* presumably feeding on a mixed diet in nature produces eggs having intermediate concentrations of MAAs (Adams & Shick 2001; Carroll & Shick, 1996). MAAs do not appear to be accumulated in mammalian cells through the diet, however, they can be absorbed by human skin cells and have the potential to serve as topical sunscreens (Mason et al., 1998).

MAAs protect marine invertebrates and phytoplankton against UV-induced damage (Adams and Shick 1996 & 2001; Neale et al., 1998). MAAs in the eggs of *S. droebachiensis*

protects them in part from UV-induced delays in cleavage (Adams and Shick, 1996) and abnormalities during later development (Adams & Shick, 2001). Fortification of eggs with MAAs may decrease cellular damage and developmental abnormalities, but the mechanism of protection remains unknown.

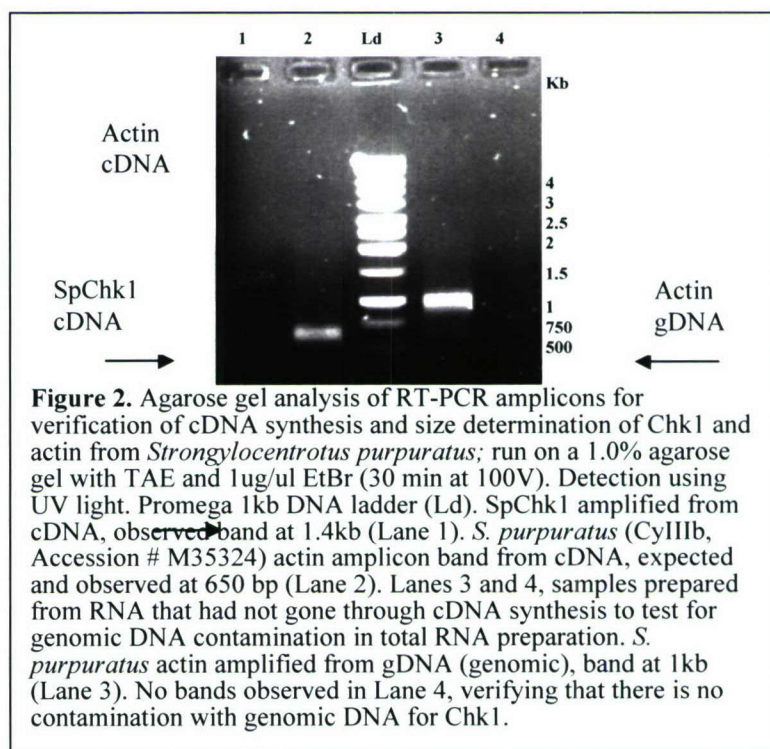
To address these aims, this year, we identified and characterized Chk1 and mycosporine-like amino acids in *S. purpuratus*. **Our specific aims were:** **Aim 1:** Identify molecular targets (markers) of radiation-induced damage that lead to perturbations in cell division using sea urchin embryos as a model system. **Aim 2-**Identify and quantify MAAs in local sea urchins, *S. purpuratus*, to create foundation for studies on the mechanism of UV-protection afforded by MAAs, use of these compounds as a molecular sensor for UV-exposure, and to search for novel MAAs.

### C. PROGRESS OF RESEARCH

**Aim 1-Identify molecular targets of UV-induced delays in cell division in sea urchin embryos, *Strongylocentrotus purpuratus* as a model.** We have examined potential targets and mechanisms of UV-induced damage in the laboratory to establish molecular markers of damage using artificial UVR that mimics the solar spectrum and levels of irradiance in coastal seawater. One of the most prominent effects of UVR on cells is disruption of the cell cycle that can lead to skin cancers and thus obvious candidates for markers of UVR-induced delay are proteins that regulate the cell cycle, including cyclin dependent kinases such as Cdc2 (Cdk1), Cdc25, and Chk1. UVR effects on the activity states of these proteins will continue to be assessed. Chk1 was the focus of this study because it is upstream of the proteins that control mitosis and it has been shown to play an important role in UV-induced damage, embryonic development and tumor suppression in other cell types (Fig.1).

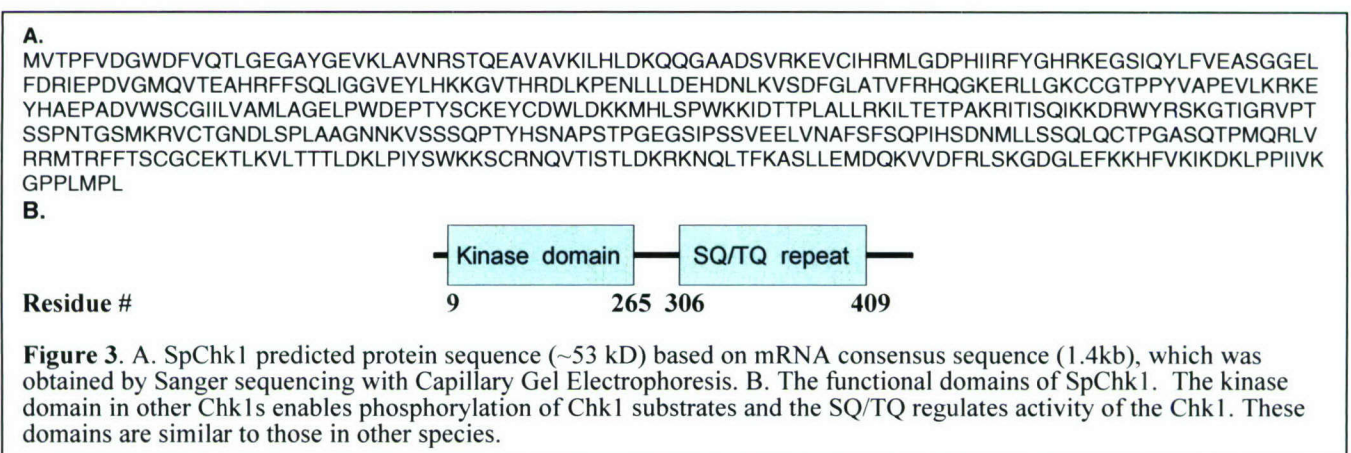
Tremendous progress has been made this past year in annotating (matching sequenced genomic regions to likely gene candidates) the *S. purpuratus* genome (Sea urchin genome Project, Human Genome Sequencing Center at the Baylor College of Medicine, [http://www.hgsc.bcm.tmc.edu/HGSC\\_home.html](http://www.hgsc.bcm.tmc.edu/HGSC_home.html)). Annotation has allowed us to better predict what cell cycle proteins are likely to occur in *S. purpuratus* eggs and embryos and has provided a foundation for us to sequence and generate antibodies that are specific for sea urchin proteins to use in functional assays (specifically when commercial antibodies have not worked). We previously demonstrated that Cdc2 activity (as indicated by dephosphorylation at tyrosine 15(Tyr 15)) is delayed in sea urchin embryos by artificial and natural UV-irradiation. We have built on these studies to examine whether Cdc25 and Chk1 (a checkpoint kinase that is activated by UV-induced DNA damage and inactivates Cdc25 in other organisms) are present and affected by UVR in sea urchin embryos. The focus of this C3RP grant was on Chk1 in sea urchins.

Our previous results using bioinformatics and western blotting indicated that Chk1 is present in sea urchin eggs and embryos



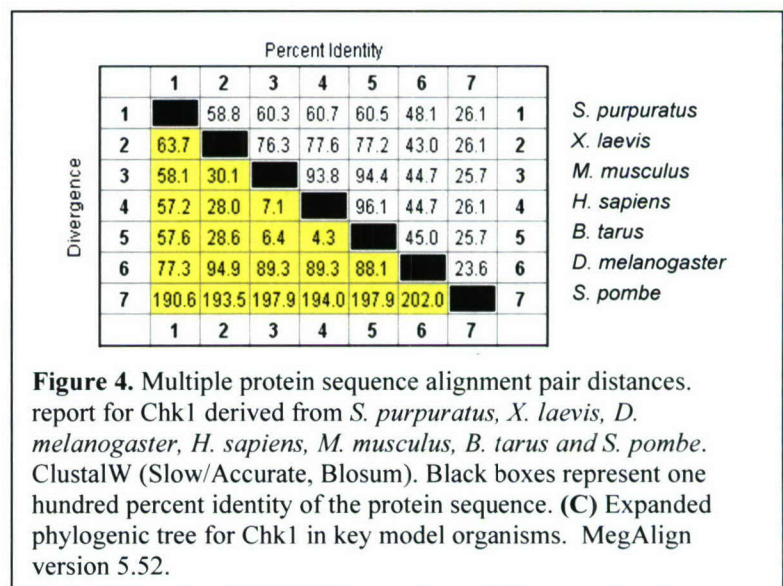
(See 2005 report). This was the first documentation of Chk1 in sea urchins. Nevertheless, even the best commercially-available antibodies for human Chk1 identified two bands in *S. purpuratus*, making it necessary to pursue better detection of Chk1. This past year, a predicted sequence for Chk1 appeared in the sea urchin genome. We personally annotated and named the Chk1 sequence in the *S. purpuratus* genome, SpChk1.

This predicted sequence allowed us to design primers, amplify, clone and sequence SpChk1 from *S. purpuratus* (Figs 2 & 3). Isolated total RNA from *S. purpuratus* eggs was donated by Dr. Foltz of U.C, Santa Barbara. cDNA was synthesized using RT-PCR with oligo-dT primers, custom Chk1 primers and total RNA. We used a positive control, CyIIIb, a *S. purpuratus* cytoskeletal actin gene that contains introns, and discriminates between actin amplified from cDNA and from gDNA. In addition, no-template RNA and RNase-treated samples controlled for amplification from contaminating nucleic acids. Fig. 2 demonstrates that we amplified cDNA for Chk1 from the total RNA lysate. The band is weak compared to the amplified actin cDNA. Nevertheless, it was abundant enough for cloning and sequencing.



The amplified Chk1 gene was cloned using the TOPO-TA Cloning Kit by Invitrogen. Chk1

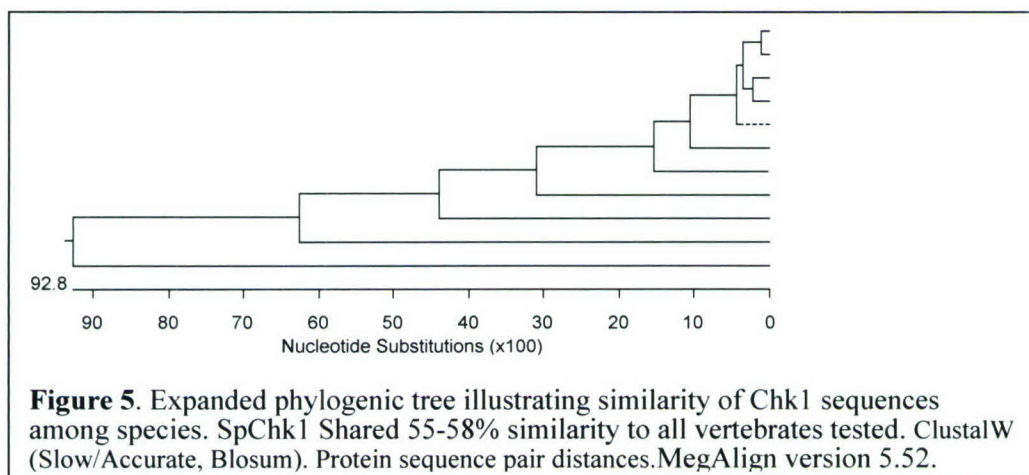
was cloned into pCR 2.1-TOPO plasmid and then chemically transformed into competent TOP10 *E. coli* cells. The remaining Chk1 PCR product was gel isolated and purified using chaotropic salts to remove small molecular weight contaminants. The transformed culture was grown overnight in LB ampicillin liquid culture as well as LB ampicillin plates with X-GAL (LBAX). The colonies grown in this medium contained the plasmid insert because it bears an ampicillin resistance gene. Five cultures were prepared from transformed plasmids on the LBAX plates. Plasmid was isolated from each transformed *E. coli* colony with Ultraclean Standard Mini Plasmid Prep Kit by MoBio. The plasmids were restriction digested with EcoR1 and Asc1 to confirm Chk1 insertion by gel electrophoresis.



The Environmental Biotechnology Laboratory (EBI), Cal Poly sequenced the isolated pCR2.1 plasmids containing the Chk1 inserts using Capillary Gel Electrophoresis. Four of the five plasmids isolated from the transformed colonies were sequenced initially with the M13 Forward and Reverse primers of pCR 2.1 vector. The inability to reliably sequence the entire 1.2kb gene required a second reaction using custom oligonucleotide primers flanking the interior region of the gene. These two sequencing reactions produced 12 sequences from which the consensus sequence was obtained. In addition, the NCBI predicted SpChk1 sequence was used as a reference for ambiguous peaks on the electropherogram. Figure 3 shows the predicted protein sequence and the predicted functional domains for SpChk1. We successfully published this *Strongylocentrotus purpuratus* SpChk1 sequence on GenBank (Accession # EF094480).

The consensus sequence was compared to existing Chk1 sequences in other organisms using BLAST and MegaAlign, DNA Star, respectively to contrast homology among several species. The phylogenetic tree shows that *Strongylocentrotus purpuratus* Chk1 diverges from mammals, but is more closely related to the mouse Chk1 sequence than any other organism (Figs. 4 & 5).

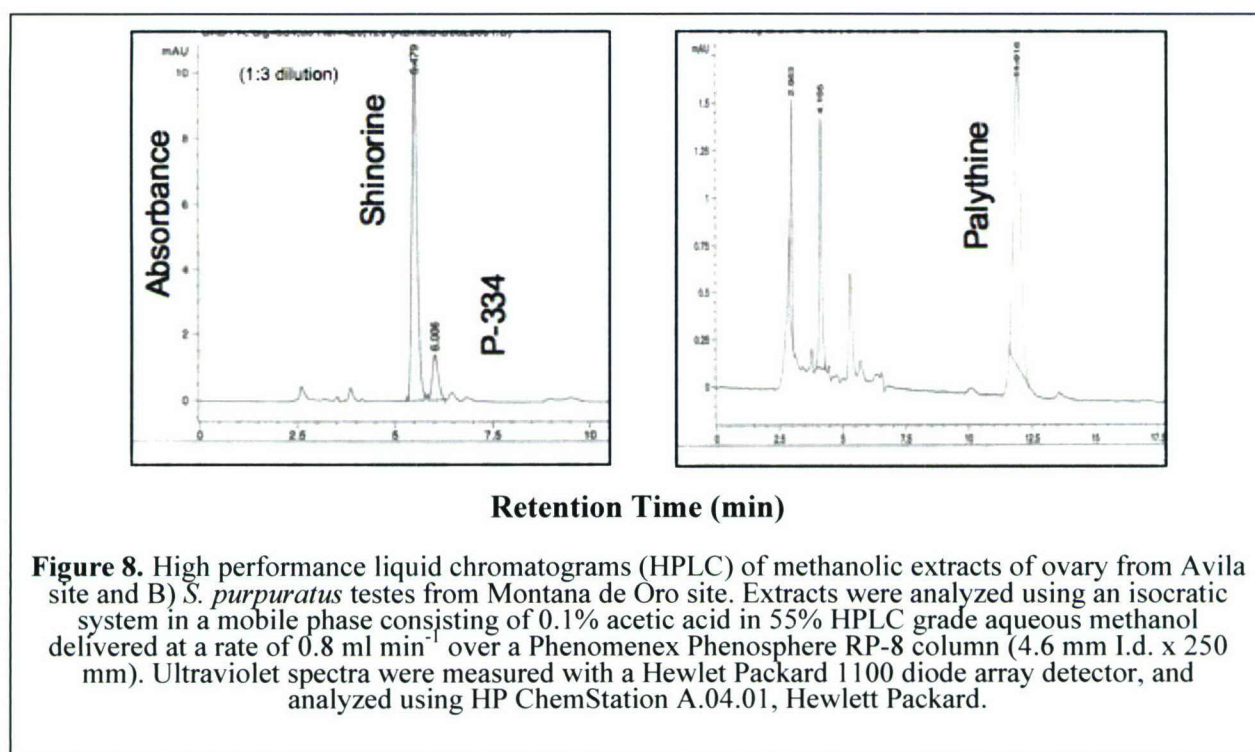
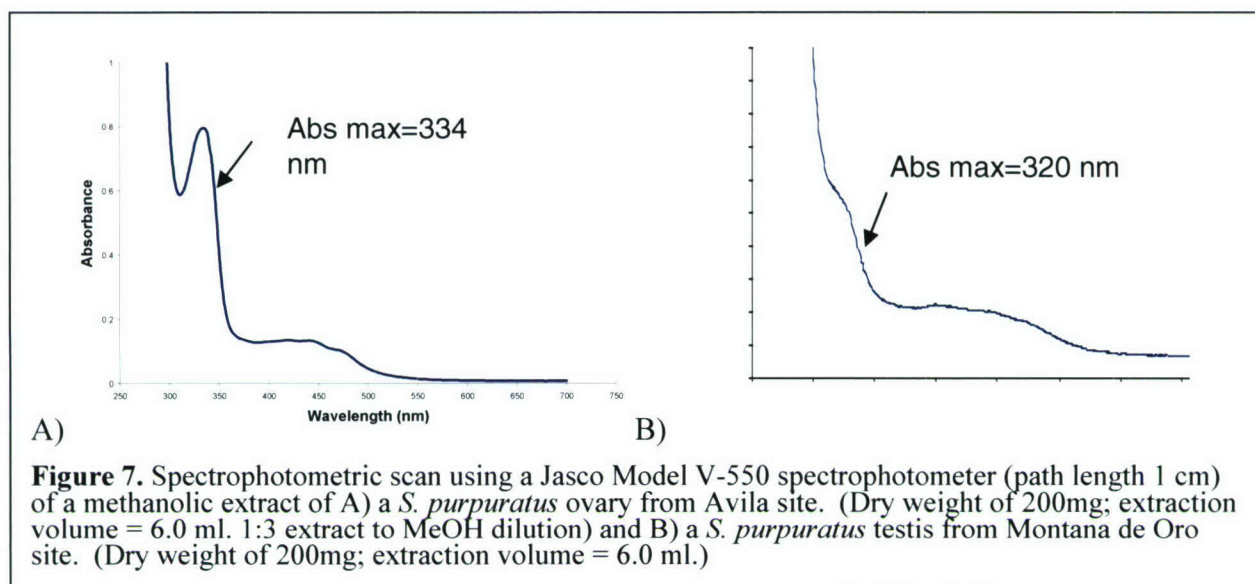
We attempted to express SpChk1 with only limited success (data not shown). Therefore, we worked with Aves Labs to produce an antibody against a peptide sequence (residues #252-268) from SpChk1. In addition, we used some of the synthetic peptide to inoculate sheep at Cal Poly as a further attempt to produce polyclonal antibodies for use in functional assays examining how UV-irradiation of sea urchin eggs and embryos affects Chk1 activity. Both antibodies successfully recognize a proteins at ~52Kd, the predicted size for SpChk1 (Fig. 6). Therefore, we are confident that we can use these antibodies in future assays examining effects of UVR on SpChk1. We will use these antibodies to examine whether Chk1 is phosphorylated at S317, S345 and S296 (indicating genotoxic stress) by UVR in sea urchin eggs and embryos.



In addition, we are currently attempting to amplify SpChk2 (another checkpoint kinase that appears in the sea urchin genome) and ATR (a protein that relays the signal of DNA damage to activate Chk1) to further characterize how components of the cell cycle signaling pathways and especially DNA damage checkpoints are affected by UVR and regulated in sea urchin embryos. These investigations may allow us to assess the entire molecular pathway between UV-induced genotoxic stress and effects of UVR on proteins that control entry into mitosis. They may also allow us to develop a molecular markers to detect UVR-induced damage in multiple marine organisms.

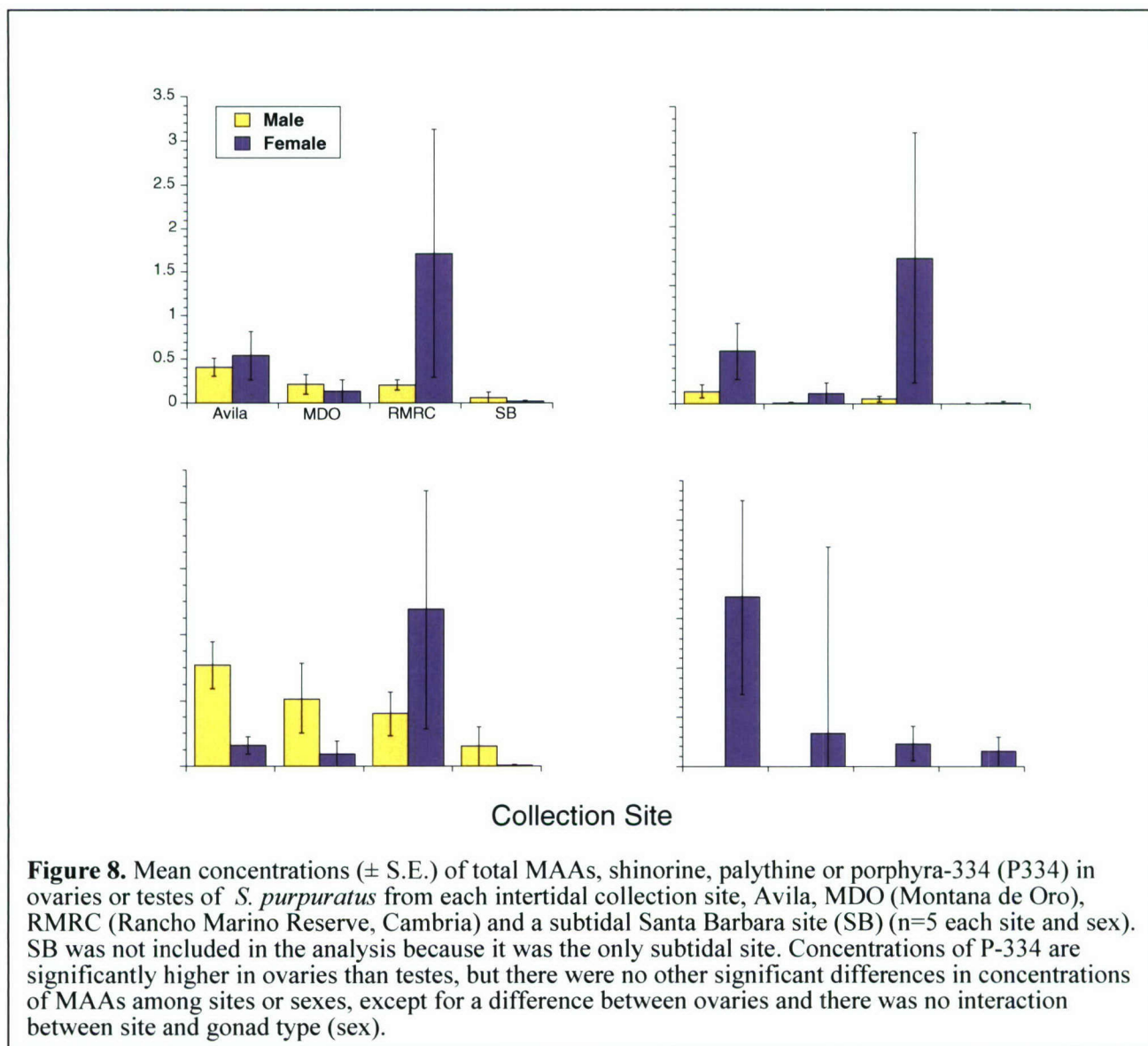
**Aim 2-Identify and quantify MAAs in local sea urchins, *S. purpuratus*, to create foundation for studies on the mechanism of UV-protection afforded by MAAs, use of these compounds as a molecular sensor for UV-exposure, and to search for novel MAAs.** My previous C<sup>3</sup>RP grant focused on characterizing and quantifying MAAs in red algae from the Central California. In addition, our previous studies have demonstrated that other adult sea urchins acquire MAAs through their diet of red algae and sequester them in their eggs and ovaries, but not in sperm. This year, we determined that MAAs are present in ovaries of some *S. purpuratus*, based on evidence from absorption spectra of methanolic extracts of these eggs.

We collected 12-15 *S. purpuratus* from each of three rocky intertidal sites of the Californian central coast within San Luis Obispo, California between late February and early March 2006 during their peak reproductive season at Hazard Beach at Montana de Oro State Park in Los Osos, San Luis Obispo Bay in Avila, and the Kenneth S. Norris Rancho Marino Reserve in Cambria. *S. purpuratus* were also collected from a subtidal site in Santa Barbara. Gonads were dissected from five adult female and adult male sea urchins and MAAs were extracted using 85% methanol.



Spectrophotometric analysis of methanolic extracts of *S. purpuratus* gonads indicate that MAAs are present in ovaries of some sea urchins (Fig. 7). We performed subsequent analysis of all samples that had an absorption peak between 309-360 nm using high performance liquid chromatography (HPLC) as per Adams et al. (2001). Three MAAs were detected in gonads of *S. purpuratus* (shinorine (absmax=334nm), porphyra-334 (P-334, absmax=334nm), and palythine (absmax=320nm)). The mean concentration for total MAAs, shinorine, P-334 and palythine in gonads was determined for each site and sex and a two-way ANOVA was performed for all intertidal sites to examine whether there was an effect of sex, site, or and interaction between site and sex (StatView, SAS Institute, V5.01). Santa Barbara samples were

not included in the analysis because this was the only subtidal site (Fig. 9). There was no significant difference in total MAA, shinorine or palythine concentrations of gonads among field sites or sexes and there was no interaction between sex and site for any MAA ( $P>0.05$ ). Nevertheless, there were generally higher concentrations of shinorine in ovaries. There were significantly higher concentrations of porphyra-334 in ovaries than testes ( $P=0.034$ ), but no difference in porphyra-334 by site ( $P>0.05$ ). Individual and total MAA concentrations of gonads were highly variable and ranged from 0 to 7.36 total MAAs nmol/mg dry wt.



These data indicate that some, but not all, *S. purpuratus* eggs contain potential photobiochemical defenses against ultraviolet radiation. Nevertheless, the concentration of MAAs in ovaries of sea urchins appear to vary within sites, indicating there may be some effect of microhabitat which vary from burrows, to exposed pool and south or north facing walls or proximity to specific types of algae that determines whether sea urchins acquire MAAs and whether embryos are protected against UVR damage.

We are currently conducting a more extensive study sampling a variety of tissues from sea urchins at more sites and sub-habitat types at each site (e.g., exposed pools, burrowing under

rocky overhang etc.) for MAAs. In addition, we will simultaneously perform algal transects on a seasonal basis to determine the availability of particular algal types, the MAA compliments in those types of algae and whether availability and MAA concentrations vary on a seasonal basis.

Overall, presence or absence of such adaptations provide an understanding of the photoadaptation strategies of *S. purpuratus* that live intertidally exposed directly to high doses of UVR and may provide further insights into how increasing UVR exposure due to decreases in stratospheric ozone levels may affect marine organisms. In addition, these studies will help us better understand mechanisms of accumulation and to identify groups of sea urchins and species of algae that we may use for studies examining how MAAs serve as sunscreens. Once we identify the patterns of MAA accumulation more conclusively we can utilize the embryos of these populations of sea urchins with variable concentrations of MAAs to examine the role of MAAs as natural sunscreens and examine whether they protect against the targets described in Aim 1.

### **Conclusion**

Together, the results of these experiments demonstrate that both Chk1 and MAAs occur in *S. purpuratus*, a promising model organism for biomedical research. This research may benefit marine science and the biotechnology industry by identifying molecular sensors of UV- damage that can be used. If our future experiments demonstrate that Chk1 is affected by UVR, we will know more about the underlying mechanisms of UV-induced delays in cell division in sea urchin embryos and we will have molecular markers to access effects of UVR on multiple organisms in the field, to assess amounts of damage occurring in nature and provide a means for testing the effectiveness of natural and commercial sunscreens.

Examination of MAAs in local algae and sea urchins will help us create a catalog of the MAAs contained in local species for future studies on dietary accumulation and protection by MAAs. In addition, it will also allow us to screen for new MAAs in a geographic area containing the most abundant array of red algae and where MAAs have not been examined previously. MAAs are presently being tested as topical sunscreens for humans. Determining the mode of protection of these compounds will help us understand their efficacy as sunscreens in many organisms, including humans.

### **USE OF FUNDS**

Funds were used to purchase materials and supplies, sequencing of Chk1, production of SpChk1 antibodies, pay for travel for students to two scientific meetings, provide release time for the PI, and provide assistantships for one graduate student, Grace Goschke (Chk1) and two undergraduate students, Laurie Gay (Chk1 cloning and sequencing) and Lindsay Chang (sea urchin MAA analysis). These students have presented these results at local and national scientific meetings noted below.

Gay, L. G. Goschke, M. Black and **N Adams** (2006). Characterization of the presence and sequence of SpChk1 from purple sea urchins, *Strongylocentrotus purpuratus*. 2006 Meeting for the American Society for Cell Biology (poster).

Chang, L and **N. Adams** (2007) Purple sea urchins, *Strongylocentrotus purpuratus*, on the Californian Central Coast contain Mycosporine-like Amino Acids (MAAs). 2007 Meeting of the Society for Integrative and Comparative Biology.

We will publish results from this work in refereed journals with students as co-authors on these papers. In addition, these data will be used to provide a foundation for future work and grant proposals. Additional funding was provided to assist in purchasing an air conditioning system for our laboratory. Our laboratory is on the south side of the Science North Building, which gets

prohibitively hot during the Spring-Fall months. This air conditioning system alleviated this problem and has allowed us to perform our protein analyses.

## REFERENCES

- Adams, NL (2001) UV radiation evokes negative phototaxis and covering behavior in the sea urchin *Strongylocentrotus droebachiensis*. *Mar Ecol Prog Ser* 213: 87-95.
- Adams, NL & Shick, JM (1996). Mycosporine-like amino acids provide protection against ultraviolet radiation in eggs of the green sea urchin, *Strongylocentrotus droebachiensis*. *Photochem and Photobiol.* 64: 149-58.
- Adams, NL and Shick, JM (2001). Mycosporine-like amino acids prevent UVB-induced abnormalities during early development of the green sea urchin *Strongylocentrotus droebachiensis*. *Mar. Biol* 138: 267-280.
- Adams, NL, Shick, JM, and Dunlap, WC (2001). Selective accumulation of mycosporine-like amino acids in ovaries of the green sea urchin, *Strongylocentrotus droebachiensis*, is not affected by UVR. *Mar. Biol.* 138: 281-294.
- Booth CR, & Morrow JH (1997) The penetration of UV into natural waters. *Photochem Photobiol* 65: 254-257
- Coleman, TR & Dunphy, WG (1994). Cdc2 regulatory factors. *Curr. Opin Cell Biol.* 6: 877-882.
- Dunlap WC, Shick JM (1998) Ultraviolet-radiation-absorbing mycosporine-like amino acids in coral reef organisms: A biochemical and environmental perspective. *J Phycol* 34: 418-430
- Fernandez-Guerra, A, Aze, J, Morales, O, Mulner-Lorillon, B, Cosson, P, Cormier, C, Bradham, N, Adams, A, Robertson, W, Marzluff, J, Coffman, and A-M Genevière. (2006). The genomic repertoire for cell cycle control and DNA metabolism in *S. purpuratus*. *Developmental Biology*. 300(1) 238-51.
- Franklin LA, Yaklovena I, Karsten U, Lüning K (1999) Synthesis of mycosporine-like amino acids in *Chondrus crispus* (Florideophyceae) and the consequences for sensitivity to ultraviolet B radiation. *J Phycol* 35: 682-693
- Gabrielli, B, Clark, J, McCormack, A, & Ellem, K (1997). Ultraviolet light-induced G2 phase cell cycle checkpoint blocks Cdc25-dependent progression into mitosis. *Oncogene*, 15: 749-58.
- Gleason DF., & Wellington, GM. (1996). Variation in UVB sensitivity of planula larvae of the coral *Agaricia agaricdes* along a depth gradient. *Mar. Biol.*, 123: 693-703.
- Häder D-P, Kumar HD, Smith RC, & Worrest RC (1998) Effects on aquatic ecosystems. In: van der Leun JC, Tang X, and Tevini M (eds), *Environmental Effects of Ozone Depletion: 1998 Assessment*, United Nations Environment Programme Report, pp 53-68
- Herzinger, T, Funk, J, Hillmer, K, Eick, D, Wolf, D, & Kind, P (1995). UVB irradiation-induced G2 cell cycle arrest in human keratinocytes by inhibitory phosphorylation of the Cdc2 cell cycle kinase. *Oncogene* 11: 2151-56.
- Jerlov, NG (1950) Ultraviolet radiation in the sea. *Nature*, 166: 111-12.
- Karentz, D. & Bosch, I (2001) Ozone-related increases in UV radiation. *Amer. Zool.* 41: 3-16.
- Karsten U, Franklin LA, Lüning K, Wiencke C (1998) Natural ultraviolet and photosynthetically active radiation induce formation of mycosporine-like amino acids in the marine macroalga *Chondrus crispus* (Rhodophyta). *Planta* 205: 257-62
- Madronich, S, McKenzie, R, Bjorn, L, & Caldwell, J. (1998). Changes in biologically active ultraviolet radiation reaching the Earth's surface. In *Environmental Effects of Ozone Depletion: 1998 Assessment* (Ed. by J. C. van der Leun, X. Tang, and M. Tevini), pp. 5-19. United Nations Environment Programme Report.
- Mason DS, Schafer F, Shick JM, & Dunlap WC (1998) Ultraviolet radiation-absorbing mycosporine-like amino acids (MAAs) are acquired from their diet by medaka fish (*Oryzias*

- latipes*) but not by SKH-1 hairless mice. *Comp Biochem Physiol* 120A: 587-598
- Meijer, L, Arion, D, Golsteyn, R, Pines, J, Brizuela, L, Hunt, T. & Beach, D (1989). Cyclin is a component of the sea urchin egg M-phase specific H1 kinase. *EMBO J.* 8:2275-2282.
- Morgan, D (1997). Cyclin-dependent kinases *Ann. rev. Cell Dev. Biol.* 13: 261-91.
- Murray A & Kirschner M (1991). What controls the cell cycle?. *Sci. Amer.* 264: 56-63.
- Neale, PJ, Banaszak, AT, & Jarriel, CR (1998) Ultraviolet sunscreens in *Gymnodinium sanguineum* (dinophyceae): Mycosporine-like amino acids protect against inhibition of photosynthesis. *J. Phycol.* 34: 928-938.
- Okano-Uchida, T, Sekiai, T, Lee, KS, Okumura, E, Tachibana, K & Kishimoto, T (1998). *In vivo* regulation of cyclin A/Cdc2 and cyclin B/Cdc2 through meiotic and early cleavage cycles in starfish. *Dev. Biol.* 197: 39-53.
- Peng, C-Y, Graves, PR, Thoma, R, Wu, Z, Shaw, A, & Piwnica-Worms, H (1997). Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25c on ser-216. *Science* 277:1501-05.
- Rusted, R (1971). Radiation response during the mitotic cycle of the sea urchin egg. In *Developmental Aspects of the Cell Cycle* (Ed. by I. Cameron, G. Padilla, & A. Zimmerman). pp. 127-159. Academic Press, New York.
- Sanchez, R., Wong, C, Thoma, R, Richman, R, Wu, Z, Piwnica-Worms, H, & Elledge, SJ(1997) Conservation of the Chk1 checkpoint in mammals:linkage of DNA damage to cdk regulation through Cdc25. *Science* 277:1497-501.
- Shick, JM, Romaine-Lioud, S, Ferrier-Pages, C, and Gattuso, J-P (1999). Ultraviolet-B radiation stimulates shikimate pathway-dependent accumulation of mycosporine-like amino acids in the coral *Stylophora pistillata* despite decreases in its population of symbiotic dinoflagellates. *Limnol. Oceanog.* 44: 1667-1682.
- Smith RC, Prezlin BB, Baker KS, Bidigare RR, Boucher NP, Coley T, Karentz D, MacIntyre S, Matlick HA, Menzies D, Ondrusek M, Wan Z, Water KJ (1992) Ozone depletion: ultraviolet radiation and phytoplankton biology in natural waters. *Science* 255: 252-259
- Smith, RC & Baker, K (1979). Penetration of UV-B and biologically effective dose-rates in natural waters. *Photochem. Photobiol.* 29: 311-323.
- Suprynowicz, FA, Groigno, L, Whitaker, M, Miller, FJ, Sluder, G, Sturrocks, J, & Whalley, T (2000) Activation of protein kinase C alters p34<sup>cdc2</sup> phosphorylation and kinase activity in early sea urchin embryos by abolishing intracellular Ca<sup>2+</sup> transients. *Biochem. J.* 349: 489-499.
- The Sea Urchin Genome Sequencing Consortium** (2006). The Genome of the Sea Urchin *Strongylocentrotus purpuratus*. *Science* 314: 941-951.
- Zhao, H and H. Piwnica-Worms (2001). ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1. *Mol Cell. Biol.* 21, 4129-4139.

## PUBLICATIONS AND PRESENTATIONS RELATED TO PROJECT

### Abstracts

Adams, Nikki and Chang, Lindsay; 'Purple sea urchins, *Strongylocentrotus purpuratus*, on the Californian Central Coast contain Mycosporine-like Amino Acids (MAAs)'; Poster presentation at the Society for Integrative and Comparative Biology, January 2007.

Mycosporine-like amino acids (MAAs) are natural sunscreens found in many marine organisms. Some marine organisms, including sea urchins, acquire MAAs through consumption of macroalgae and sequester them in their gonads and gametes. MAAs are produced in red algae in San Luis Obispo, CA and are potentially eaten by local sea urchins. The aim of this study was to determine whether purple sea urchins, *Strongylocentrotus purpuratus* on the Californian central coast have MAAs in their ovaries or testes. We compared concentrations of MAAs in gonads of

sea urchins from three rocky intertidal sites using high performance liquid chromatography of methanolic extracts of gonads. Three specific MAAs (shinorine, porphyra-334 and palythine) are present in ovaries and testes of *S. purpuratus*, but there was no significant difference in total MAA, shinorine (absmax=334nm) or palythine (absmax=320nm) concentrations of gonads among field sites or sexes and there was no interaction between sex and site for any MAA ( $P>0.05$ ). Nevertheless, there were generally higher concentrations of shinorine in ovaries. There were significantly higher concentrations of porphyra-334 (absmax=334nm) in ovaries ( $P=0.034$ ), but no difference in porphyra-334 by site. Individual and total MAA concentrations of gonads were highly variable and ranged from 0 to 7.36 nmol/mg dry wt., indicating that there may be some effect of microhabitat or proximity to specific types of algae that determines whether sea urchins acquire MAAs and whether embryos are protected against UVR damage. These data also indicate that some, but not all, *S. purpuratus* eggs contain potential photobiochemical defenses against ultraviolet radiation.

*Adams, Nikki, et al.; 'Identification and Sequence of Checkpoint Kinase 1 from Purple Sea Urchins, Strongylocentrotus purpuratus'; Poster Presentation at the American Society for Cell Biology Meeting, December 2006.*

Exposure to ultraviolet radiation (UVR) causes developmental delays, abnormalities and decreased survivorship in the sea urchin *Strongylocentrotus purpuratus*. Mammalian cells maintain internal checkpoints that arrest the cell cycle at both the G1/S and G2/M transitions if mutations are detected. Although these checkpoints have not previously been characterized in sea urchins, studies by our laboratory demonstrate a Cdc2-dependent developmental delay of embryos upon exposure to UVR. We are studying members of the protein signal transduction pathway that regulate the cell cycle via Cdc2, most specifically, Checkpoint Kinase 1 (Chk1). It is predicted that UVR will activate sea urchin Chk1, as in other model organisms, through upstream regulators such as ATM/ATR. This increased Chk1 activity would result in the inhibition of Cdc25, a phosphatase responsible for driving the G2/M transition via activation of Cdc2. Recently, the *S. purpuratus* genome was sequenced and is being annotated through the Baylor College Human Genome Sequencing Center. Based upon the characterization of Chk1 orthologs in humans and several other organisms, a predicted sea urchin Chk1 sequence was identified. This sequence was used to design primers that successfully amplified Chk1 cDNA derived from sea urchin egg mRNA. Overlapping sequences were obtained from several cloned isolates and the resulting consensus was found to contain an additional exon that had not been included in the original predicted sequence. A synthetic peptide was created based upon this sequence and is currently being used to raise urchin-specific anti-Chk1 antibodies for further analysis of the role that UVR plays in Chk1 activation.

### **Full Publications**

Copies of full publications follow.



**Fig. 1.** Purple sea urchins (*S. purpuratus*) grazing on the remains of a giant kelp hold fast after an unusually heavy recruitment in Carmel Bay, California (7). [Photo courtesy of J. M. Watanabe]

There are important differences among the different species: Some are more effective as grazers than others, and they vary in their diets,

growth rates, longevity, and importance in fisheries. Some show no sign of senility and live for well over a century (17). Elucidation of their

genomes will open new avenues of research into the underlying genetic and evolutionary bases of these variations.

#### References and Notes

1. C. Harrold, J. S. Pearce, *Echinoderm Studies* **2**, 137 (1987).
2. L. Rogers-Bennett, in *Edible Sea Urchins: Biology and Ecology*, J. M. Lawrence, Ed. (Elsevier, Amsterdam, Netherlands, 2007), pp. 393–425.
3. P. K. Dayton *et al.*, *Ecol. Monogr.* **62**, 421 (1992).
4. A. W. Ebeling *et al.*, *Mar. Biol.* **84**, 287 (1985).
5. K. D. Lafferty, *Ecol. Appl.* **14**, 1566 (2004).
6. J. S. Pearce, A. H. Hines, *Mar. Ecol. Prog. Ser.* **39**, 275 (1987).
7. J. M. Watanabe, C. Harrold, *Mar. Ecol. Prog. Ser.* **71**, 125 (1991).
8. T. A. Ebert *et al.*, *Mar. Ecol. Prog. Ser.* **111**, 41 (1994).
9. J. B. Jackson *et al.*, *Science* **293**, 629 (2001).
10. P. K. Dayton *et al.*, *Ecol. Appl.* **8**, 309 (1998).
11. J. A. Estes, D. O. Duggins, *Ecol. Monogr.* **65**, 75 (1995).
12. M. H. Graham, *Ecosystems* **7**, 341 (2004).
13. M. S. Foster, *Hydrobiologia* **192**, 21 (1990).
14. M. J. Tegner, P. K. Dayton, *ICES J. Mar. Sci.* **57**, 579 (2000).
15. L. Rogers-Bennett, J. S. Pearce, *Conserv. Biol.* **15**, 642 (2001).
16. D. Sweetnam *et al.*, *Calif. Coop. Oceanic Fish. Invest. Rep.* **46**, 10 (2005).
17. T. A. Ebert, J. R. Southon, *Fish. Bull.* **101**, 915 (2003).
18. I thank J. Watanabe for providing the photo used in Fig. 1, L. Rogers-Bennett for sharing her manuscript with me, V. Pearce and an anonymous reviewer for providing comments on the manuscript, and R. A. Cameron for inviting me to prepare it.

10.1126/science.1131888

#### RESEARCH ARTICLE

## The Genome of the Sea Urchin *Strongylocentrotus purpuratus*

Sea Urchin Genome Sequencing Consortium\*†

We report the sequence and analysis of the 814-megabase genome of the sea urchin *Strongylocentrotus purpuratus*, a model for developmental and systems biology. The sequencing strategy combined whole-genome shotgun and bacterial artificial chromosome (BAC) sequences. This use of BAC clones, aided by a pooling strategy, overcame difficulties associated with high heterozygosity of the genome. The genome encodes about 23,300 genes, including many previously thought to be vertebrate innovations or known only outside the deuterostomes. This echinoderm genome provides an evolutionary outgroup for the chordates and yields insights into the evolution of deuterostomes.

The genome of the sea urchin was sequenced primarily because of the remarkable usefulness of the echinoderm embryo as a research model system for modern molecular, evolutionary, and cell biology. The sea urchin is the first animal with a sequenced genome that (i) is a free-living, motile marine invertebrate; (ii) has a bilaterally organized embryo but a radial adult body plan; (iii) has the endoskeleton and water vascular system found only in echinoderms; and (iv) has a nonadaptive immune

system that is unique in the enormous complexity of its receptor repertoire. Sea urchins are remarkably long-lived with life spans of *Strongylocentrotus* species extending to over a century [see supporting online material (SOM)] and highly fecund, producing millions of gametes each year; and *Strongylocentrotus purpuratus* is a pivotal component of subtidal marine ecology and an important fishery catch in several areas of the world, including the United States. Although a research model in developmental biology for a

century and a half, for most of that time, few were aware of one of the most important characteristics of sea urchins, a character that directly enhances its significance for genomic analysis: Echinoderms (and their sister phylum, the hemichordates) are the closest known relatives of the chordates (Fig. 1 and SOM). A description of the echinoderm body plan, as well as aspects of the life-style, longevity, polymorphic gene pool, and characteristics that make the sea urchin so valuable as a research organism, are presented in the SOM.

The last common ancestors of the deuterostomal groups at the branch points shown in Fig. 1 are of Precambrian antiquity [>540 million years ago (Ma)], according to protein molecular phylogeny. Stem group echinoderms appear in the Lower Cambrian fossil assemblages dating to 520 Ma. Cambrian echinoderms came in many distinct forms, but from their first appearance, the fossil record illustrates certain distinctive features that are still present: their water vascular system, including rows of tube feet protruding through holes in the ambulacral grooves and their calcitic endoskeleton (mainly, a certain form of  $\text{CaCO}_3$ ).

\*Correspondence should be addressed to George M. Weinstock. E-mail: gwstock@bcm.tmc.edu

†All authors with their contributions and affiliations appear at the end of this paper.

which displays the specific three-dimensional structure known as "stereom." The species sequenced, *Strongylocentrotus purpuratus*, commonly known as the "California purple sea urchin" is a representative of the thin-spined "modern" group of regularly developing sea urchins (euechinoids). These evolved to become the dominant echinoid form after the great Permian-Triassic extinction 250 million years ago.

We present here a description of the *S. purpuratus* genome and gene products. The genome provides a wealth of discoveries about the biology of the sea urchin, Echinodermata, and the deuterostomes. Among the key findings are the following:

- The sea urchin is estimated to have 23,300 genes with representatives of nearly all vertebrate gene families, although often the families are not as large as in vertebrates.
- Some genes thought to be vertebrate-specific were found in the sea urchin (deuterostome-specific); others were identified in sea urchin but not the chordate lineage, which suggests loss in the vertebrates.
- Expansion of some gene families occurred apparently independently in the sea urchin and vertebrates.
- The sea urchin has a diverse and sophisticated immune system mediated by an astonishingly large repertoire of innate pathogen recognition proteins.
- An extensive defenseome was identified.
- The sea urchin has orthologs of genes associated with vision, hearing, balance, and chemosensation in vertebrates, which suggests hitherto unknown sensory capabilities.
- Distinct genes for biomineralization exist in the sea urchin and vertebrates.
- Orthologs of many human disease-associated genes were found in the sea urchin.

## Sequencing and Annotation of the *S. purpuratus* Genome

**Sequencing and assembly:** Sperm from a single male was used to prepare DNA for all libraries (tables S1 and S2) and whole-genome shotgun (WGS) sequencing. The overall approach was based on the "combined strategy" used for the rat genome (1), where WGS sequencing to six times coverage was combined with two times sequence coverage of BAC clones from a minimal tiling path (MTP) (fig. S1). The use of BACs provided a framework for localizing the assembly process, which aided in the assembly of repeated sequences and solved problems associated with the high heterozygosity of the sea urchin genome, without our resorting to extremely high coverage sequencing.

Several different assemblies were produced during the course of the project (see SOM for details). The Sea Urchin Genome Project (SUGP) was the first to produce both intermediate WGS assemblies and a final combined assembly. This was especially useful, not only for the early

availability of an assembly for analysis, but also because WGS contigs were used to fill gaps between BACs in the combined assembly. The pure WGS assembly was produced (v 0.5 GenBank accession number range AAGJ01000001 to AAGJ01320773; also referred to as NCBI build 1.1) and released in April 2005. The final combined BAC-WGS assembly was released in July 2006 as version (v) 2.1 and submitted to GenBank (accession number range AAGJ02000001 to AAGJ02220581).

A second innovation in the SUGP was the use of the clone-array pooled shotgun sequencing (CAPSS) strategy (2) for BAC sequencing (fig. S2). The MTP consisted of 8248 BACs, and rather than prepare separate random libraries from each of these, the CAPSS strategy involved BAC shotgun sequencing from pools of clones and then deconvoluting the reads to the individual BACs. This allowed the BAC sequencing to be performed in 1/5th the time and at 1/10th the cost.

The principal new challenge in the SUGP was the high heterozygosity in the outbred animal that was sequenced. It was known that single-copy DNA in the sea urchin varied by as much as 4 to 5% [single nucleotide polymorphism (SNP) plus insertion/deletion (indel)], which is much greater than human (~0.5%) (3). Moreover, alignment of WGS reads to the early v 0.1 WGS assembly revealed at least one SNP per 100 bases, as well as a comparable frequency of indel variants. This average frequency of a mismatch per 50 bases or higher prevented merging by the assembly module in Atlas, the Phrap assembler, and also made it difficult to determine if reads were from duplicated but diverged sections of the genome or heterozygous homologs. This challenge was met by adding components to Atlas to handle local regions of heterozygosity and to take advantage of the BAC data, because each BAC sequence represented a single haplotype (see SOM). High heterozygosity has been seen in the past with the *Ciona* genomes (4, 5) and is likely to be the norm in the future as fewer inbred organisms are sequenced. Moreover, the CAPSS approach makes BAC sequencing more manageable for large genomes. Thus, the sea urchin project may serve as a paradigm for future difficult endeavors.

Combining the BAC-derived sequence with the WGS sequence generated a high-quality draft with 4 to 5% redundancy that covered more than 90% of the genome while sequencing to a level of 8× base coverage (table S2). The assembly size of 814 Mb is in good agreement with the previous estimate of genome size, 800 Mb ± 5% (6). The assembly is a mosaic of the two haplotypes, but it was possible to determine the phase of the BACs on the basis of how many mismatches neighboring BACs had in their overlap regions. This information will be used to create a future version of the genome in which the individual haplotypes are resolved.

**Gene predictions.** The v 0.5 WGS assembly displayed sufficient sequence continuity (a contig

N50 of 9.1 kb) and higher-order organization (a scaffold N50 of 65.6 kb) to allow gene predictions to be produced and the annotation process to begin even while the BAC component was being sequenced. We generated an official gene set (OGS), consisting of ~28,900 gene models, by merging four different sets of gene predictions with the GLEAN program (7) (see SOM for details). One of these gene sets, produced from the Ensembl gene prediction software, was created for both v 0.5 and v 2.0 assemblies.

To estimate the number of genes in the *S. purpuratus* genome, we began with the 28,900 gene models in the OGS and reduced this by the 5% redundancy found by mapping to the v 2.0 assembly, then increased it by a few percent for the new genes observed in the Ensembl set from the v 2.0 assembly compared with v 0.5. From manual analysis of well-characterized gene sets (e.g., ciliary, cell cycle control, and RNA metabolism genes), we estimated that, in addition to redundancy, another 25% of the genes in the OGS were fragments, pseudogenes, or otherwise not valid. Finally, whole-genome tiling microarray analysis (see below) showed 10% of the transcriptionally active regions (long open reading frames, not small RNAs) were not represented by genes in the OGS. Taken together, this analysis gave an estimate of about 23,300 genes for *S. purpuratus*. Information on all annotated genes can be found at (8).

The overall trends in gene structure were similar to those seen in the human genome. The statistics of the Ensembl predictions from the WGS assembly revealed an average of 8.3 exons and 7.3 introns per transcript (see SOM). The average gene length was 7.7 kb with an average primary transcript length of 8.9 kb. A broad distribution of all exon lengths peaked at around 100 to 115 nucleotides, whereas that for introns at around 750 nucleotides. The smaller average intron size relative to humans' was consistent with the trend that intron size is correlated with genome size.

**Annotation process.** Manual annotation and analysis of the OGS was performed by a group of over 200 international volunteers, primarily from the sea urchin research community. To facilitate and to centralize the annotation efforts, an annotation database and a shared Web browser, Genboree (9), were established at the BCM-HGSC. These tools enabled integrated and collaborative analysis of both precomputed and experimental information (see SOM). A variety of precomputed information for each predicted gene model was made available to the annotators in the browser, including expressed sequence tag (EST) data, the four unmerged gene prediction sets, and transcription data from whole-genome tiling microarray with embryonic RNA (see below) (10). Additional resources available to the community are listed in table S4.

Over 9000 gene models were manually curated by the consortium with 159 novel models (gene models not represented in the OGS) added to the official set. If we assume no bias in the curated gene

models, the number of novel models added may imply that the official set contains >98% of the protein-coding genes.

**Genome features.** A window on the genetic landscape is scaffold-centric in *S. purpuratus*, because linkage and cytogenetic maps are not available. The 36.9% GC content of the genome is uniformly low because assessment of the average GC content by domains is consistent (36.8%), and the distribution is tight (see SOM). Genes from the OGS show no tendency to occupy regions of higher- or lower-than-average GC content. In fact, nearly all genes lie in regions of 35 to 39% GC.

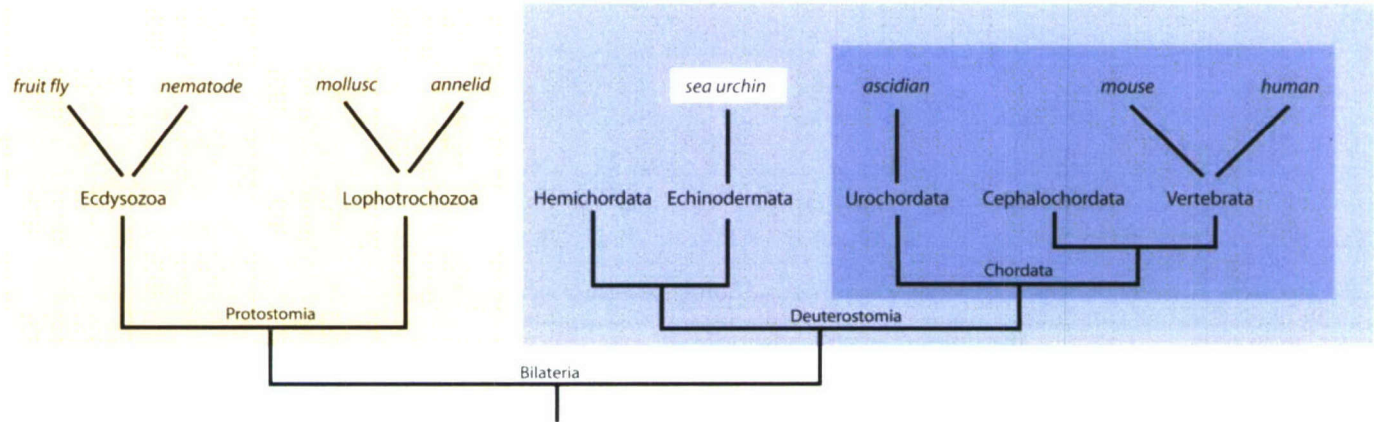
The Echinoderm Genome in the Context of Metazoan Evolution

The sea urchin genetic tool kit lends evolutionary perspective to the gene catalogs that characterize the superclades of the bilaterian animals. The distribution of highly conserved protein domains and sequence motifs provides a view of the expansion and contraction of gene families, as well as an insight into changes in protein function. Examples are enumerated in Table 1, which presents a global overview of gene variety

obtained by comparing sequences identified in Interpro, and Table 2, which shows the distribution of specific Pfam database domains associated with selected aspects of cell physiology, including sequences identified in the cnidarian *Nematostella vectensis* (11). The Interpro data suggest that about one-third of the 50 most prevalent domains in the sea urchin gene models are not in the 50 most abundant families in the other representative genomes (mouse, tunicate, fruit fly, and nematode), and thus, they constitute expansions that are specific at least to sea urchins, if not to the complex of echinoderms and hemichordates. Two of the most abundant domains make up 3% of the total and mark genes that are involved in the innate immune response. Others define proteins associated with apoptosis and cell death regulation, as well as proteins that serve as downstream effectors in the Toll-interleukin 1 (IL-1) receptor (TIR) cascade. The quinoprotein amine dehydrogenase domain seen in the sea urchin set is 10 times as abundant as in other representative genomes and may be used in the systems of quinone-containing pigments known to occur in these marine animals. The large number of nu-

cleosomal histone domains found agrees with the long-established sea urchin-specific expansion of histone genes. In summary, the distribution of proteins among these conserved families shows the trend of expansion and shrinkage of the preexisting protein families, rather than frequent gene innovation or loss. Gene family sizes in the sea urchin are more closely correlated with what is seen in deuterostomes than what is seen in the protostomes.

Of equal interest are the sorts of proteins not found in sea urchins. The sea urchin gene set shares with other bilaterian gene models about 4000 domains, whereas 1375 domains from other bilaterian genomes are not found in the sea urchin set. In agreement with the lack of morphological evidence of gap junctions in sea urchins, there are no gap junction proteins (connexins, pannexins, and innexins). Also missing are several protein domains unique to insects, such as insect cuticle protein, chitin-binding protein, and several pheromone- or odorant-binding proteins, as well as a vertebrate invention—the Krüppel-associated box or KRAB domain, a repressor domain in zinc finger transcription factors (12). Finally, searches for specific subfamilies of G protein-coupled



**Fig. 1.** The phylogenetic position of the sea urchin relative to other model systems and humans. The chordates are shown on the darker blue background overlapping the deuterostomes as a whole on a lighter blue background. Organisms for which genome projects have been initiated or finished are shown across the top.

**Table 1.** Unique aspects of gene family distribution in sea urchin: Selected examples of the frequency of Interpro domains in the proteome of selected species. ID is the identification number used in the INTERPRO database; the second column shows the name given to the domain or motif family in the database. Species abbreviations: Sp, *Strongylocentrotus purpuratus*; Mm, *Mus musculus*; Ci, *Ciona intestinalis*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*.

ID	Name	Species, total number (percentage of total matches)				
		Sp	Mm	Ci	Dm	Ce
IPR001190	Speract/scavenger receptor	361 (1.79)	14 (0.08)	1 (0.01)	2 (0.02)	0 (0.00)
IPR000157	TIR	248 (1.23)	22 (0.12)	9 (0.09)	9 (0.09)	2 (0.02)
IPR011029	DEATH-like	172 (0.85)	8 (0.05)	19 (0.18)	5 (0.05)	1 (0.01)
IPR007111	NACHT nucleoside triphosphatase	135 (0.67)	16 (0.09)	28 (0.27)	0 (0.00)	0 (0.00)
IPR011044	Quinoprotein amine dehydrogenase, $\beta$ chain-like	122 (0.60)	7 (0.04)	15 (0.15)	5 (0.05)	6 (0.05)
IPR000558	Histone H2B	110 (0.54)	14 (0.08)	2 (0.02)	100 (1.00)	17 (0.13)
IPR001951	Histone H4	93 (0.46)	7 (0.04)	0 (0.00)	101 (1.01)	16 (0.12)
IPR002119	Histone H2A	87 (0.43)	24 (0.14)	2 (0.02)	104 (1.04)	19 (0.14)
IPR008042	Retrotransposon, Pao	76 (0.38)	0 (0.00)	0 (0.00)	0 (0.00)	6 (0.05)
IPR000164	Histone H3	72 (0.36)	17 (0.10)	5 (0.05)	103 (1.03)	22 (0.17)

receptors (GPCRs) that are known as chemo-sensory and/or odorant receptors in distinct bilaterian phyla failed to detect clear representatives in the sea urchin genome. However, this failure more likely reflects the independent evolution of these receptors, rather than a lack of chemoreceptive molecules, because the sea urchin genome encodes close to 900 GPCRs of the same superfamily (rhodopsin-type GPCRs), several of which are expressed in sensory structures (13). A conservative way to compare gene sets is to count the strict orthologs that give reciprocal BLAST matches. Genes that are genuine orthologs are likely to yield each other as a best hit. Comparison of sea urchin, fruit fly, nematode, ascidian, mouse, and human gene sets (Fig. 2) indicates that the greatest number of reciprocal best matches is observed between mouse and human, which reflects their close rela-

tion. The numbers of presumed orthologous genes between the ascidian and the two mammals are about equal, but are less than the number counted between these species and the sea urchin. The difference is consistent with the lower gene number and reduced genome size in the urochordates (4). The number of reciprocal pairs for sea urchin and mouse is about 1.5 times the matches between proteins in sea urchin and fruit fly. The number of nematode proteins matching either sea urchin or fruit fly is even lower. This is likely the result of the more rapid sequence changes in the nematode compared with the other species used in this analysis. More than 75% of the genes that are shared by sea urchin and fruit fly are also shared between sea urchin and mouse. Thus, these genes constitute a set of genes common to the bilaterians,

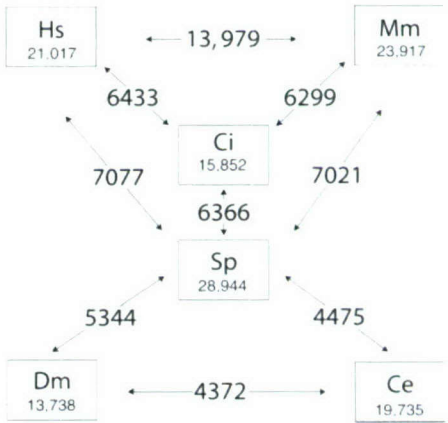
whereas the additional sea urchin–mouse pairs are unique to the deuterostomes. The sea urchin genome consequently provides evidence for the now extremely robust concept of the deuterostome superclade. A 1908 concept that originated in the form of embryos of dissimilar species (14) is demonstrated by genomic comparisons. Developmental Genomics In the 1980s, the sea urchin embryo became the focus of cis-regulatory analyses of embryonic gene expression, and there was a great expansion of molecular explorations of the developmental cell biology, signaling interactions, and regulatory control systems of the embryo. Analysis of the entire genome facilitated the first large-scale correlation of the gene regulatory network for development, which represents the genomic control circuitry for specification of the endoderm and mesoderm of this embryo (15–17) with the encoded potential of the sea urchin. The embryo transcriptome and regulome. Because of indirect development in the sea urchin, embryogenesis is clearly separated from adult body plan formation, in developmental process and in time, and therefore, it is possible to estimate the genetic repertoire specifically required for formation of a simple embryo (10). Pooled mRNA preparations from four stages of development, up to the mid-late gastrula stage (48 hours), were hybridized with a whole-genome tiling array. Expression of about 12,000 to 13,000 genes, as conservatively assessed, was seen during this early period, indicating that ~52% of the entire protein-

**Table 2.** Distribution among sequenced animal genomes of various Pfam domains associated with selected aspects of eukaryotic cell physiology. In *S. purpuratus*, the number of annotated genes is listed; the number in parentheses is the total number of models (including ones that were not annotated) predicted to contain the Pfam domain. For *Nematostella vectensis* (Nv), numbers were obtained by searching Stellabase (11).

Process	Domain	PFAM no.	Sp	Hs	Dm	Ce	Nv
Cell cycle control	Cyclin_N	PF00134	15 (17)	21	11	7	7
	Cyclin_C	PF02984	7 (8)	12	4	5	4
	E2F_TDP	PF02319	3 (5)	11	3	4	3
	RB_A	PF01858	2	3	2	1	0
	RB_B	PF01857	2	3	2	1	0
	P53	PF00870	1	3	1	1	0
	Cullin	PF00888	7	9	8	7	4
Histone metabolism	Skp1	PF01466	1	3	5	21	1
	Histone*	PF00125	49	75	8	?	?
	Linker histone*	PF00538	5	8	2	?	?
	Nucleo-plasmin	PF03066	2	5	2	0	1
	NAP	PF00956	2	24	4	2	0
	HDAC	PF00850	8	11	5	8	3
	DOT1	PF08123	1	1	1	6	1
RNA metabolism	RRM_1	PF00076	140 (178)	245	126	99	41
	TUDOR	PF00567	15	13	15	8	7
	DEAD	PF00270	93 (125)	78	56	65	27
	LSM	PF01423	17	21	17	18	4
	KH-1	PF00013	28 (31)	36	28	28	5
	DSRM	PF00035	14 (15)	21	14	13	8
	3'-5'-Exo-nuclease	PF01612	13 (15)	5	5	9	5
Apoptosis†	Exonuc_X-T	PF00929	9 (11)	15	7	10	5
	Caspase	PF00656	31 (33)	14	7	4	5
	BIR	PF00653	4 (7)	8	4	1	4
	Bcl-2	PF00452	10	11	2	1	7
	TNFR_c6	PF00020	8 (9)	8	1 (no DD)	1 (no DD)	2
	NACHT	PF05729	129 (145)	18	1	1	2
	NB-ARC	PF00931	3	1	1	1	0
	DEATH	PF00531	47 (101)	30	9	6	6
	DED	PF01335	4 (5)	7	1	0	5
	CARD	PF00619	5 (10)	20	1	0	8

- Complexity intermediate between that in vertebrates and protostome invertebrate model organisms
- Complexity greater than that found in other model organisms
- Complexity lower than that found in other model organisms

\*Numbers of histone genes refer to distinct core or linker histone genes, as opposed to total gene number as a result of large tandemly repeated arrays (e.g., ~400 clusters of early histone arrays in sea urchin, 100 copies of a tandem array in *Drosophila*, with each array containing a gene for the four core and one H1 histone). †Numbers for Hs, Dm, and Ce obtained from (53).



**Fig. 2.** Orthologs among the Bilateria. The number of 1:1 orthologs captured by BLAST alignments at a match value of  $e = 1 \times 10^{-6}$  in comparisons of sequenced genomes among the Bilateria. The number of orthologs is indicated in the boxes along the arrows, and the total number of International Protein Index database sequences is shown under the species symbol. Hs, *Homo sapiens*; Mm, *Mus musculus*; Ci, *Ciona intestinalis*; Sp, *S. purpuratus*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*.

coding capacity of the sea urchin genome is expressed during development to the mid-late gastrula stage. An additional set of microarray experiments extended the interrogation of embryonic expression to the 3-day pluteus larva stage (see SOM) (18).

The DNA binding domains of transcription factor families are conserved across the Bilateria, and these protein domain motifs were used to extract the sea urchin homologs (see SOM). For each identified gene, if data were not already available, probes were built from the genome sequence and used to measure transcript concentration by quantitative polymerase chain reaction with a time series of embryo mRNAs, as well as to determine spatial expression by whole-mount *in situ* hybridization.

All bilaterian transcription factor families were represented in the sea urchin with a few rare exceptions (see below), so the sea urchin data strongly substantiate the concept of a pan-bilaterian regulatory tool kit (19) or “regulome.” We found that 80% of the whole sea urchin regulome (except the zinc finger genes) was expressed by 48 hours of embryogenesis (20), an even greater genetic investment than the 52% total gene use in the same embryo.

**Signal transduction pathways.** More than 1200 genes involved in signal transduction were identified. Comparative analysis highlights include the protein kinases that mediate the majority of signaling and coordination of complex pathways in eukaryotes. The *S. purpuratus* genome has 353 protein kinases, intermediate between the core vertebrate set of 510 and the fruit fly and nematode conserved sets of ~230. Fine-scale classification and comparison with annotated kinomes (21, 22) reveals a remarkable parsimony. Indeed, with only 68% of the total number of human kinases, the sea urchin has members of 97% of the human kinase subfamilies, lacking just four of those subfamilies (Axl, FastK, H11, and NKF3), whereas *Drosophila* lacks 20 and nematodes 32 (Fig. 3) (23). Most sea urchin kinase subfamilies have just a single member, although many are expanded in vertebrates; thus, the sea urchin kinome is largely nonredundant. The sea urchin therefore possesses a kinase diversity surprisingly comparable to that of vertebrates without the complexity. A small number of kinases were more similar to insect than to vertebrate homologs (including the Titin homolog Projection, the Syk-like tyrosine kinase Shark, and several guanylate cyclases), which indicated for the first time the loss of kinase classes in vertebrates (23). Expression profiling showed that 87% of the signaling kinases and 80% of the 91 phosphatases were expressed in the embryo (23, 24), which emphasized the importance of signaling pathways in embryonic development.

The small guanosine triphosphatases (GTPases) function as molecular switches in signal transduction, nuclear import and export, lipid metabolism, and vesicle docking. Vertebrate

GTPase families were expanded after their divergence from echinoderms, in part by whole-genome duplications (25–27). The sea urchin genome did not undergo a whole-genome duplication, yet phylogenies for four Ras GTPase families (Ras, Rho, Rab, and Arf) revealed that local gene duplications occurred (Fig. 4), which ultimately resulted in a comparable number of monomeric GTPases in the human and sea urchin genomes (28). Thus, expansion of each family in vertebrates and echinoderms was achieved by distinct mechanisms (gene-specific versus whole genome duplication). More than 90% of the small GTPases are expressed during sea urchin embryogenesis, which suggests that the complexity of signaling through GTPases is comparable between sea urchins and vertebrates.

The Wnt family of secreted signaling molecules plays a central role in specification and patterning during embryonic development. Phylogenetic analyses from cnidarian to human indicate that of the 13 known Wnt subfamilies, *S. purpuratus* has 11, missing Wnt2 and Wnt11 homologs (Fig. 5). *S. purpuratus* has WntA, previously reported as being absent from deuterostomes (29). Of 126 genes described as components of the Wnt signal transduction machinery, homologs of ~90% were present in the sea urchin genome, which indicates a high level of conservation of all three Wnt pathways (30). However, of 94 Wnt transcriptional target genes reported in the literature, mostly from vertebrates (31), only 53% were found with high confidence in the sea urchin genome (Fig. 6). The absent Wnt targets include vertebrate adhesion molecules, which were frequently missing from the sea urchin genome (32), as well as signaling receptors, which are more divergent and thus more difficult to identify. In contrast, most transcription factor targets of the Wnt pathway are present in the genome, which reflects a higher degree of conservation of transcription factor families (20). Taken together, the genomic analysis of signal transduction components indicates that sea urchins have signaling machinery strikingly comparable to that of vertebrates, often without the complexity that arises from genetic redundancy.

### Sea Urchin Biology

Analysis of the genome allows understanding of parts of the organism that have not been well studied. Several examples of this follow with further details in the SOM. Additional areas such as intermediary metabolism, metalloproteases, ciliary structure, fertilization, and germline specification are presented in the SOM.

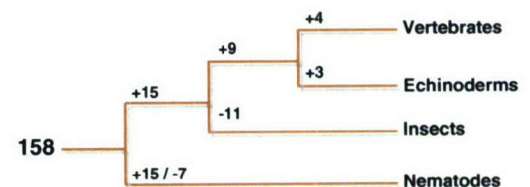
### Defense Systems

The need to deal with physical, chemical, and biological challenges in the environment underlies the evolution of an array of defense gene families and pathways. One set of protective mechanisms involves the

immune system, which responds to biotic stressors such as pathogens. A second group of genes comprises a chemical “defensome,” a network of stress-sensing transcription factors and defense proteins that transform and eliminate many potentially toxic chemicals.

**The sea urchin immune system.** The sea urchin has a greatly expanded innate immunity repertoire compared with any other animal studied to date (table S5). Three classes of innate receptor proteins are particularly increased (Fig. 7). These make up a vast family of Toll-like receptors (TLRs), a similarly large family of genes that encode NACHT and leucine-rich repeat (LRR)-containing proteins (NLRs), and a set of genes encoding multiple scavenger receptor cysteine-rich (SRCR) domain proteins of a class highly expressed in the sea urchin immune cells or coelomocytes (33, 34). Receptors from each of these families participate in immunity by recognizing nonself molecules that are conserved in pathogens or by responding to self molecules that indicate the presence of infection (35). In contrast, homologs of signal transduction proteins and nuclear factor kappa B (NFκB)/Rel domain transcription factors that are known to function further downstream of these genes were present in numbers similar to those in other invertebrate species. One of the more unexpected findings from our analysis of sea urchin immune genes was the identification of a Rag1/2-like gene cluster (36). The presence of this cluster, along with other recent findings (37), suggested the possibility that these genes had been part of animal genomes for longer than previously considered. Further analysis of the genomic insights into the innate immune system and the underpinnings of vertebrate adaptive immunity can be found in a review in this issue (38).

**The complement system.** The complement system of vertebrates is a complex array of soluble serum proteins and cellular receptors arranged into three activation pathways (classical, lectin, and alternative) that converge and activate the terminal or lytic pathway. This system opsonizes pathogenic cells for phagocytosis and sometimes activates the terminal pathway, which leads to pathogen destruction. An invertebrate complement system was first identified in the sea urchin [for reviews, see (39, 40)], and



**Fig. 3.** Protein kinase evolution: Invention and loss of protein kinase subfamilies in metazoan lineages. Deuterostomes share 9 protein kinase subfamilies absent from *C. elegans* and *Drosophila*, and the sea urchin has not lost any of the 158 metazoan primordial kinase classes, unlike insects or nematodes. [From (23)]

## Sea Urchin Genome

the analysis of the genome sequence presented a more complete picture of this important immune effector system. In chordates, collectins initiate the lectin cascade through members of the mannose-binding protein (MBP)-associated protease (MASP)/C1r/C1s family. Several genes encoding collectins, C1q and MBP, have been predicted (39) and were present in the genome; however, members of the MASP/C1r/C1s family were not identified. There was no evidence for the classical pathway, which links the complement cascade with immunoglobulin recognition in jawed vertebrates. The alternative pathway is initiated by members of the thioester protein family, which, in the sea urchin, was somewhat expanded with four genes. Two of the thioester proteins, SpC3 and SpC3-2, are known to be expressed, respectively, in coelomocytes and in embryos and larvae. Furthermore, there were three homologs of factor B, the second member of the alternative pathway (41).

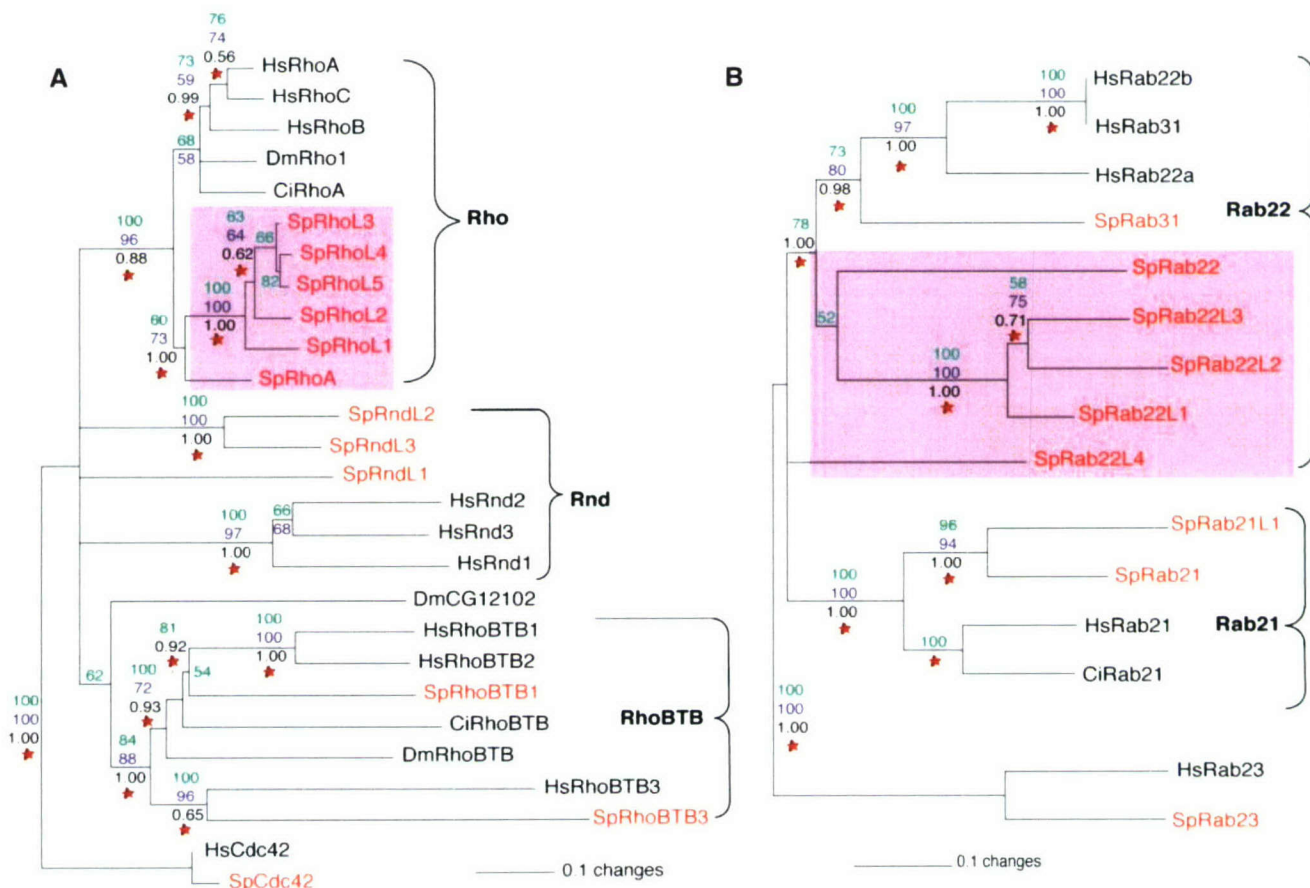
The terminal complement pathway in vertebrates acts to destroy pathogens or pathogen-infected cells with large pores called membrane attack complexes (MACs). Twenty-eight gene models were identified that encode MAC-

perforin domains, but none of these had the additional domains expected for terminal complement factors (C6 through C9). Instead, these are members of a novel and very interesting gene family with perforin-like structure. In vertebrates, perforins carry out cell-killing functions by cytotoxic lymphocytes through the formation of small pores in the cell membranes. If the complement system in the sea urchin functions through multiple lectin and alternative pathways in the absence of the lytic functions of the terminal pathway, the major activity of this system is expected to be opsonization.

**Homologs of immune regulatory proteins.** Cytokines are key regulators of intercellular communication involving immune cells, acting to coordinate vertebrate immune systems. Genes encoding cytokines and their receptors often evolve at a rapid pace, and most families are known only from vertebrate systems. Although members of many cytokine, chemokine, and receptor families were not identified in the sea urchin genome, a number of important immune signaling homologs were present. These included members of the tumor necrosis factor (TNF) ligand and receptor superfamilies, an IL-1 receptor and accessory

proteins, two IL-17 receptor-like genes and 30 IL-17 family ligands, and nine macrophage inhibitory factor (MIF)-like genes. Receptor tyrosine kinases (RTKs) included those that bind important growth factors that regulate cell proliferation in vertebrate hematopoietic systems. Of particular note, from the sea urchin genome, were two vascular endothelial growth factor (VEGF) receptor-like genes and a Tie1/2 receptor, all of which were expressed in adult coelomocytes. Many of these genes are homologs of important inflammatory regulators and growth factors in higher vertebrates, and these sea urchin homologs may have similar functions in regulating coelomocyte differentiation and recruitment.

Representatives of nearly all subclasses of important vertebrate hematopoietic and immune transcription factors were present in the sea urchin genome. Notably, the genome contained homologs of immune transcription factors that had not been identified previously outside of chordates, including PU.1/SpiB/SpiC, a member of the Ets subfamily, and a zinc finger gene with similarity to the Ikaros subfamily. Transcript prevalence measurements showed that PU.1, the Ikaros-like gene and homologs of Gata1/2/3, E2A/HEB/ITF2, and Stem

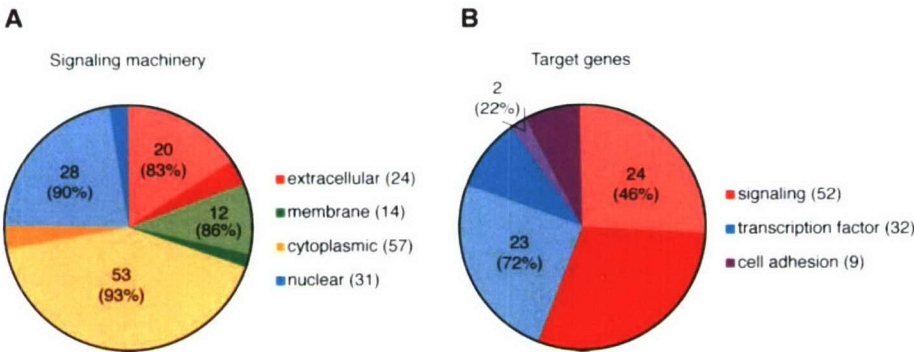


**Fig. 4.** Partial phylogenies of the Rho (A) and the Rab families (B) of small GTPases. The pink boxes highlight gene-specific duplications that increased sea urchin GTPase numbers, resulting in a complexity comparable to vertebrates. Numbers at each

junction represent confidence values obtained via three independent phylogenetic methods [neighbor-joining (green), maximum parsimony (blue), and Bayesian (black)]; red stars indicate nodes retained by maximum likelihood. [From (28)]

Families	Genes	Distribution of the Activators of the Wnt signals throughout the animal kingdom							
		Cnidarians	Ecdysozoans	Lophotrochozoans	Echinoderms	Urochordates	Cephalochordates	Vertebrates	
		Sea anemone ( <i>Nematostella vectensis</i> )	Fly ( <i>Drosophila melanogaster</i> )	Annelid and Molluscs ( <i>Platynereis dumerilii</i> and <i>Patella vulgata</i> )	Sea Urchin ( <i>Strongylocentrotus purpuratus</i> )	Ascidian ( <i>Ciona intestinalis</i> )	Amphioxus ( <i>Branchiostoma floridae</i> )	Human ( <i>Homo sapiens</i> )	
	wnt1					?			
	wnt2/13		X		X				
	wnt3		X						
	wnt4		X			?			
	wnt5			( <i>D. japonica</i> )					
	wnt6								
	wnt7								
	wnt8		X			X			
	wnt9/14/15								
	wnt10								
	wnt11		X		X	X			
	wnt16		( <i>C. elegans-egl20</i> )			?			
	wntA		( <i>A. gambiae</i> )			X		X	

**Fig. 5.** Survey of the Wnt family of secreted signaling molecules in selected metazoans. Each square indicates a single Wnt gene identified either through genome analyses or independent studies, and squares with a question mark indicate uncertainty of the orthology. Letter X's represent absence of members of that subfamily in the corresponding annotated genome; empty spaces have been left for species for which genomic databases are not yet available. [From (30)]



**Fig. 6.** Presence of Wnt signaling machinery components (A) and target genes (B) in the *S. purpuratus* genome. (A) The 126 genes involved in the transduction of the Wnt signals have been separated into four categories from the extracellular compartment to the nucleus. Sea urchin homologs are identified by the lighter shade (indicated by both the number and the percentage of homologs that were identified within the chart); the total number of known genes is indicated in the chart legend. (B) The 93 reported Wnt targets have been divided into three categories: signaling molecules, transcription factors, and cell adhesion molecules. Colors and numbers are as in (A).

Cell Leukemia (SCL) were all expressed at substantial levels in coelomocytes (41). This was consistent with the presence of conserved mechanisms of regulating gene expression among sea urchin coelomocytes and vertebrate blood cells.

**ABC transporters.** Many chemicals are removed from cells by efflux proteins known as ATP-binding cassette (ABC) or multidrug efflux transporters. *S. purpuratus* has 65 ABC transporter genes in the eight major subfamilies of these genes [ABC A to H; (42)]. The ABCC family of multidrug transporters is about 25% larger than in other deuterostome genomes with at least 30 genes in this family (nearly half of the sea urchin ABC transporters), and 25 of these 30 genes showed substantial mRNA expression in eggs, embryos, or larvae. Much of the expansion is in the Sp-ABCC5 and Sp-

ABCC9 families, whereas orthologs of the vertebrate gene ABCC2 (also called MRP2) are absent. Because the ABCC family is known to generally transport more hydrophilic compounds than other transporter families, such as the ABCB genes, sea urchins may have increased need for transport of these compounds. ABCC efflux activity has been described in sea urchin embryos and, consistent with the genomic expansion of the ABCC family, the major activity in early embryos ensues from an ABCC-like efflux mechanism.

**Cytochrome P-450 monooxygenase (CYP).** Enzymes in the CYP1, CYP2, CYP3, and CYP4 families carry out oxidative biotransformation of chemicals to more hydrophilic products. The sea urchin has 120 CYP genes, and those related to CYP gene families 1 to 4 constitute 80% of the total,

which suggests that there has been selective pressure to expand functionality in these gene families (42). Eleven CYP1-like genes are present in the sea urchin genome, more than twice the number in chordates. CYP2-like and CYP3-like genes are also present at greater numbers than in other deuterostomes. In addition to the CYPs in families 1 to 4, the sea urchin genome contains homologs of proteins involved in developmental patterning (CYP26), cholesterol synthesis (CYP51), and metabolism (CYP27, CYP46). Homologs of some CYPs with endogenous functions in vertebrates were not found; however, (CYP19, androgen aromatase; CYP8, prostacyclin synthase; CYP11, pregnenolone synthase; CYP7, cholesterol-7 $\alpha$ -hydroxylase). These CYP genes in concert with additional expanded defensive gene families represent a large diversification of defense gene families by the sea urchin relative to mammals (42).

**Oxidative defense and metal-complexing proteins.** The metal-complexing proteins include three metallothionein genes and three homologs of phytochelatin synthase genes. Genes for antioxidant proteins include three superoxide dismutase (SOD) genes and a gene encoding ovoperoxidase (an unusual peroxidase with SOD-like activity), along with one catalase, four glutathione peroxidase, and at least three thioredoxin peroxidase genes. Reactive oxygen detoxification genes may be important in conferring the long life-span of sea urchins, because oxidative damage is thought to be a major factor in aging.

**Diversity and conservation in xenobiotic signaling.** The diversity of genes encoding xenobiotic-sensing transcription factors that regulate biotransformation enzymes and transporters was similar to other invertebrate genomes, but in most cases lower than vertebrates.

# Sea Urchin Genome

For example, the sea urchin genome encoded a single predicted CNC-bZIP protein homologous to the four human CNC-bZIP proteins involved in the response to oxidative stress. There were two sea urchin homologs of the aryl hydrocarbon receptor (AHR), which in vertebrates mediates the transcriptional response to polynuclear and halogenated aromatic hydrocarbons and, in both protostomes and deuterostomes, also regulates specific developmental processes (43–45). One of the sea urchin AHR homologs was more closely related to the vertebrate AHR; the other shared greatest sequence identity with the *Drosophila* AHR homolog *spineless*. Sea urchins also had two genes encoding hypoxia-inducible factors (HIF $\alpha$  subunits), which regulate adaptive responses to hypoxia, and a gene encoding ARNT, a PAS protein that is a dimerization partner for both AHRs and HIFs.

*Strongylocentrotus purpuratus* has 32 nuclear receptor (NR) genes (20), two-thirds the number in humans, including several with potential roles in chemical defense (42). The sea urchin genome also contains two peroxisome proliferator-activated receptor (PPAR, NR1C) homologs and an NR1H gene coorthologous to both liver X receptor (LXR) and farnesoid X receptor (FXR) (42). Genes homologous to the vertebrate xenobiotic sensor NR11 genes [pregnane X receptor, PXR; constitutive androstane receptor, CAR (46)] are absent, although three NR1H-related genes

were found, which possibly form a new subfamily of genes involved in xenobiotic sensing.

Many of the defense genes are expressed during development (10, 42), which suggests that they have dual roles in chemical defense and in developmental signaling. In several cases (CYPs, AHR, NF-E2), the evolution of pathways for chemical defense may have involved recruitment from developmental signaling pathways (42).

## Nervous System

The echinoderm nervous system is the least well studied of all the major metazoan phyla. For a number of technical reasons, the structure and function of echinoderm nerves have been neglected. Analysis of the sea urchin genome has enabled an unprecedented glimpse into the neural and sensory functions and has revealed several novel molecular approaches to the study of echinoderm nervous systems (Table 3).

The nervous systems of echinoderm larvae and adults are dispersed, but they are not simple nerve nets. This organization differs from both vertebrates, which do not have a dispersed nervous system, and hemichordates, which do have nerve nets (47). Adult sea urchins have thousands of appendages, each with sensory neurons, ganglia, and motor neurons arranged in local reflex arcs. These peripheral appendages are connected to each other and to radial nerves, which provide overall control and coordination (47, 48).

Nearly all of the genes encoding known neurogenic transcription factors are present in the sea urchin genome, and several are expressed in neurogenic domains before gastrulation, which indicates that they may operate near the top of a conserved neural gene regulatory network (47). Axon guidance molecules known from other metazoans are also expressed in the developing embryo. Unexpectedly, genes encoding the neurotrophin-Trk receptor system that were thought to be vertebrate-specific because they were not found in *Ciona*, are present in sea urchin, which suggests a deuterostome origin and a potential loss in urochordates.

The genes required to construct neurons and to transmit signals are present, but the repertoire of neural genes and the initial characterization of expression of a number of them led to unex-

pected and surprising conclusions. There appear to be no genes encoding gap junction proteins, which suggests that communication among neurons depends on chemical synapses without ionic coupling. The repertoire of sea urchin neurotransmitters is large, but melatonin and adrenalin are lacking, as they are in ascidians (4, 47). Cannabinoid, lysophospholipid, and melanocortin receptors are not present in urchins, but orthologs were found in ascidians (4, 47). In contrast, some sets of genes thought to be chordate-specific have sea urchin orthologs, for example, insulin and insulin-like growth factors (IGFs) that are more similar to their chordate counterparts than those of other invertebrates (47). Overall, the genome contains representatives of all five large superfamilies of GPCRs, including those that mediate signals from neuropeptides and peptide hormones. Both the secretin and rhodopsin superfamilies display marked lineage-specific expansions (13, 47).

**Sensory systems.** There were 200 to 700 putative chemosensory genes that formed large clusters and lacked introns, which are features of chemosensory genes in vertebrates, but not in *Caenorhabditis elegans* and *Drosophila melanogaster*. Many of these genes encoded amino acid motifs that were characteristic of vertebrate chemosensory and odorant receptors (13, 47). Sea urchins had an elaborate collection of photoreceptor genes that quite surprisingly appeared to be expressed in tube feet (13, 47). These included many genes encoding transcription factors regulating retinal development and a photorhodopsin gene.

Human Usher syndromes are genetic diseases affecting hearing, balance, and retinitis pigmentosa (retinal photoreceptor degeneration). Most of the genes involved have been identified, and they encode a set of membrane and cytoskeletal proteins that form an interacting network that controls the arrangement of mechanosensory stereocilia in hair cells of the mammalian ear. Many or all of the proteins play some roles in photoreceptor organization and/or maintenance. Orthologs of virtually the entire set of membrane and cytoskeletal proteins of the Usher syndrome network were found in the sea urchin genome. These include the very large membrane proteins, usherin and VLGR-1 and large cadherins (Cadh23 and possibly Pcad15), all of which participate in forming links between stereocilia in mammalian hair cells, as well as myosin 7 and 15, two PDZ proteins (harmonin and whirlin) and another adaptor protein (SANS), which participate in linking these membrane proteins to the cytoskeleton. In addition, two membrane transporters, NBC (a candidate Usher syndrome target known to interact with harmonin) and TrpA1 (the mechanosensory channel connected to the tip links containing cadherin 23), have orthologs in the sea urchin genome. Sea urchins do not have ears or eyes, so they must deploy these proteins in

	TLR	NLR	SRCR	PGRP	GNBP	C3/4/5	Bf/C2	C1q/MBP	Terminal pathway
<i>H.s.</i> 10 (+1 $\psi$ )	20	81 (16)	6	0	3	2	5	+	
<i>C.i.</i>	3	0	22 (8)	6	0	2	3	11	+/-
<i>S.p.</i>	222	203	1095 (218)	5	3	4	3	8	-
<i>D.m.</i>	9	0	14 (7)	15	4	0*	0	0	-
<i>C.e.</i>	1	0	3 (1)	0	0	0	0	0	-

**Fig. 7.** Gene families encoding important innate immune receptors and complement factors in animals with sequenced genomes. For some key receptor classes, gene numbers in the sea urchin exceeds other animals by more than an order of magnitude. Representative animals include *H.s.*, *Homo sapiens*; *C.i.*, *Ciona intestinalis*; *S.p.* *Strongylocentrotus purpuratus*; *D.m.* *Drosophila melanogaster*; and *C.e.* *Caenorhabditis elegans*. Indicated gene families include TLR, toll-like receptors; NLR, NACHT and leucine-rich repeat (LRR) domain-containing proteins similar to the vertebrate Nod/NALP genes; SRCR, Scavenger receptor cysteine-rich domain genes; PGRP, peptidoglycan recognition protein domain genes; and GNBP, Gram-negative binding proteins. C3/4/5, thioester proteins homologous to vertebrate C3, C4, and C5; Bf/C2, complement factors homologous to vertebrate C2 and factor B; C1q/MBP, homologs of vertebrate lectin pathway receptors; and Terminal pathway, homologs of vertebrate C6, C7, C8, and C9. SRCR gene statistics are given as domain number/gene number for multiple SRCR-containing proteins (numbers for *C. intestinalis* includes all SRCR proteins). Asterisk in the *D. melanogaster* C3/4/5 column is meant to denote the presence of related thioester genes (TEPs) and a true C3/4/5 homolog from another arthropod. +/- for *C. intestinalis* Terminal pathway column indicates the presence of genes with similarity to C6 only (Nonaka and Yoshizaki 2004). Phylogenetic relations among species are indicated by a cladogram at the left.

other sensory processes. Sea urchins respond to light, touch, and displacement and probably use some of same sensory genes used by vertebrates.

### The Echinoderm Adhesome

The *S. purpuratus* genome contained representatives of all the standard metazoan adhesion receptors (table S7), but the emphasis on different classes of receptors differed substantially from that used by vertebrates. The integrin family was intermediate in size between those of protostomes and vertebrates—several chordate-specific expansions of the integrin repertoire were absent, and there were some expansions unique (so far) to echinoderms. The cadherin repertoire was also small relative to vertebrates (a dozen or so instead of over a hundred), and many chordate-specific expansions were missing. Specialized large cadherins shared by protostomes and vertebrates were present, as well as some specialized large cadherins previously thought to be chordate-specific, but overall, the cadherin repertoire was more invertebrate than vertebrate in character. Sea urchins lacked the integrins and cadherins

that link to intermediate filaments in vertebrates.

In contrast, sea urchins had large repertoires of adhesion molecules containing immunoglobulin superfamily, fibronectin type 3 repeat (FN3), epidermal growth factor (EGF), and LRR repeats. In addition to the expansion of TLRs and NLRs mentioned above, there are large expansions of other LRR receptor families, including GPCRs (32). The key neural adhesion systems involved in regulating axonal outgrowth were present (netrin/Unc5/DCC; Slit/Robo; and semaphorins/plexins), as were adhesion molecules involved in synaptogenesis (Agrin/MUSK; and neuroligin/neurexins). This was not surprising because these molecules were known in both protostomes and vertebrates. However, structurally, the synapses of echinoderms are unusual because there are no direct synaptic contacts (49). Some of them were expressed in sea urchin embryos before there are any neurons, suggesting that they may have other roles as well.

The basic metazoan basement membrane extracellular matrix (ECM) tool kit was present—

two alpha-IV collagen genes, perlecan, laminin subunits, nidogen, and collagen XV/XVIII. There did not appear to be much, if any, expansion of these gene families, as is found in vertebrates, which suggests that there is less diversity among basement membranes. Quite a few ECM proteins present in chordates, but not protostomes, were also missing in sea urchins, including fibronectins, tenascins, von Willebrand factor, vitronectin, most vertebrate-type matrix proteoglycans, and complex VWA/FN3 collagens among others (32). Absence of these genes may be related to the absences of neural crest migration, a high shear endothelial-lined vasculature and, of course, cartilage and bone.

In addition to the components of Usher syndromes mentioned above, it was surprising to find a clear ortholog of reelin, a large ECM protein involved in establishing the layered organization of neurons in the vertebrate cerebral cortex. Reelin is mutated in the *reeler* mouse, and mutations in the *reeler* gene in humans have been associated with Norman-Roberts-type lissencephaly syndrome. Reelin has a unique domain

**Table 3.** Genomic insights into sea urchin neurobiology.

Neural process	Revelations from the genome	Genes
Neural development	Neurogenic ectoderm is specified in early embryonic development.	Sp-Achaete-scute, Sp-homeobrain, Sp-Rx (retinal anterior homeobox), Sp-Zic2
Synapse structure and function	Echinoderm synapses are structurally unusual, despite the presence of many genes encoding proteins involved in synapse function.	Sp-Neurologgin, Sp-neurexin, Sp-agrin, Sp-MUSK, Sp-thrombospondin, Sp-Rim2, Sp-Rab3, exocyst complex, Snares, SM, synaptotagmins
Electrical signaling and coupling	Neurons have ion channel proteins, but lack electrical coupling via gap junctions.	Voltage-gated K <sup>+</sup> , Ca <sup>2+</sup> , and Na <sup>+</sup> channels, but no connexins or pannexins/innexins
Neurotransmitter/neuromodulatory diversity	Neurons use the same neurotransmitters as vertebrates, but lack melatonin and adrenalin.	Enzymes involved in synthesis, transport, reception, and hydrolysis of serotonin, dopamine, noradrenaline, $\gamma$ -aminobutyric acid (GABA), histamine, acetylcholine, glycine, and nitric oxide
GPCR signaling	Identification of GPCRs that are unique to chordates and identification of expanded GPCR families.	Orthologs of vertebrate cannabinoid, lysophospholipid, and melanocortin receptors are absent; 162 secretin receptor-like genes
Peptide signaling	G protein-coupled peptide receptors indicate diversity in peptide signaling systems, but only a few sea urchin neuropeptides or peptide hormones identified.	37 G protein-coupled peptide receptors. Precursors for SALMFamides, NGFFFamide, and a vasotocin-like peptide
Neurotrophins	Neurotrophins and neurotrophin receptors are not unique to chordates.	Sp-Neurotrophin, Sp-Trk, Sp-p75NTR, ependymins
Insulin and IGFs	More similar to vertebrate forms than invertebrate insulin-like molecules.	Sp-IGF1, Sp-IGF2
Chemosensory functions	A large family of predicted chemoreceptor genes, some expressed in tube feet or pedicellariae, indicates a complex chemosensory system.	Over 600 genes encoding putative G protein-coupled chemoreceptors, many tandemly repeated and lacking introns
Photoreception functions	Genes associated with photoreception are expressed in tube feet.	Photorhodopsins, Sp-Pax6, retinal transcription factors
Mechanosensory functions	Orthologs of vertebrate mechanosensory genes are present.	Sp-Usherin, Sp-VLGR-1, Sp-cadherins, Sp-myosin 7, Sp-myosin 15, Sp-harmonin, Sp-whirlin, Sp-NBC, Sp-TrpA1

composition and organization (Reeler, EGF, BNR) that has not been found outside chordates, but the sea urchin genome included a very good homolog of reelin. Receptors for reelin are believed to include low-density lipoprotein receptor-related proteins (LRPs), and there are a number of these receptors in *S. purpuratus* although it is as yet unclear whether they are reelin receptors, lipoprotein receptors, or something else. Similar receptors are also involved in human disease (atherosclerosis).

## Biom mineralization Genes

Among the deuterostomes, only echinoderms and vertebrates produce extensive skeletons. The possible evolutionary relations between biomineralization processes in these two groups have been controversial. Analysis of the *S. purpuratus* genome revealed major differences in the proteins that mediate biomineralization in echinoderms and vertebrates (50). First, there were few sea urchin counterparts of extracellular proteins that mediate biomineral deposition in vertebrates. For example, in vertebrates, an important class of proteins involved in biomineralization is the family of secreted, calcium-binding phosphoproteins, or SCPPs. Sea urchins did not have counterparts of SCPP genes, which supports the hypothesis that this family arose via a series of gene duplications after the echinoderm-chordate divergence (51). Second, almost all of the proteins that have been directly implicated in the control of biomineralization in sea urchins were specific to that clade. The echinoderm skeleton consists of magnesium calcite (as distinct from the calcium phosphate skeletons of vertebrates) in which is occluded many secreted matrix proteins. The sea urchin spicule matrix proteins were encoded by a family of 16 genes that are organized in small clusters and likely proliferated by gene duplication. Counterparts of sea urchin spicule matrix genes were not found in vertebrates, amphioxus, or ascidians. Likewise, other genes that have been implicated in biomineralization in sea urchins, including genes that encode the transmembrane protein P16 and MSP130, a glycosylphosphatidylinositol-linked glycoprotein, were members of small clusters of closely related genes without apparent homologs in other deuterostomes. The members of all three of these sea urchin-specific gene families were expressed specifically by the biomineral-forming cells of the embryo, the primary mesenchyme cells [see (50)]. As a whole, these findings highlighted substantial differences in the primary sequences of the proteins that mediate biomineralization in echinoderms and vertebrates.

**Cytoskeletal genes.** In addition to identifying genes for all previously known *S. purpuratus* actins and tubulins, one  $\delta$ - and two  $\epsilon$ -tubulin genes were found (52). Newly identified motor protein genes include members of four more classes of myosin, and eight more families of kinesins.

The first dynein cloned and sequenced was from sea urchin, and although most *S. purpuratus* dynein heavy chain genes mapped one-to-one to mammalian homologs, Sp-DNAH9 mapped one-to-three, as it was equidistant between the closely similar mammalian genes DNAH9, DNAH11, and DNAH17 (52).

## Conclusions

Our estimate of 23,300 genes is similar to estimates for vertebrates, despite the fact that two whole-genome duplications are believed to have occurred in the chordate lineage after divergence from the lineage leading to the echinoderms (25–27). From the analysis presented here, it seems likely that many mechanisms shaped the final genetic content of these genomes. On the one hand, there are cases of gene families that are expanded in vertebrates compared with sea urchin, including examples of the expected 4:1 ratio from two duplications (15). However other patterns are also found. The nuclear receptor family is only slightly reduced in sea urchin compared with that of humans, which suggests gene loss followed the vertebrate duplications. The unprecedented expansions of innate immune system diversity contrast sharply with the much smaller sets of counterparts that are present in the sequenced genomes of protostomes, *Ciona*, and vertebrates, an example of independent expansion in the sea urchin, whereas the GTPases described here have expanded in sea urchin to about the same numbers as in vertebrates. Thus, whereas the duplications of the chordate lineage were a contributor to the increased complexity of vertebrates, regional expansions clearly play a large role in the evolution of these animals.

The refinement of the inventory of vertebrate-specific or protostome-specific genes likewise benefits from the sea urchin genome. Many more human genes have shared ancestry across the deuterostomes, and in fact, bilaterian genes are more broadly shared than had been inferred from comparison of the previously limited genome sequences. The new biological niche sampled by the sea urchin genome provides not only a clearer view of the deuterostome and bilaterian ancestor, but has also provided a number of surprises. The finding of sea urchin homologs for sensory proteins related to vision and hearing in humans may lead to interesting new concepts of perception, and the extraordinary organization of the sea urchin immune system is different from any animal yet studied. From a practical standpoint, the sea urchin may be a treasure trove. Because of the many pathways shared by sea urchin and human, the sea urchin genome includes a large number of human disease gene orthologs. Many of the genes described in the preceding sections fall into this category (see tables S8 and S9) and cover a surprising diversity of systems such as nervous, endocrine, and blood systems, as well as

muscle and skeleton, as exemplified by the Huntington and muscular dystrophy genes. Continued exploration of the sea urchin immune system is expected to uncover additional variations for protection against pathogens. The immense diversity of pathogen-binding motifs encoded in the sea urchin genome provides an invaluable resource for antimicrobial applications and the identification of new deuterostome immune functions with direct relevance to human health. These exciting possibilities show that much biodiversity is yet to be uncovered by sampling additional evolutionary branches of the tree of life.

## References and Notes

1. R. A. Gibbs et al., *Nature* **428**, 493 (2004).
2. W. W. Cai, R. Chen, R. A. Gibbs, A. Bradley, *Genome Res.* **11**, 1619 (2001).
3. R. J. Britten, A. Cetta, E. H. Davidson, *Cell* **15**, 1175 (1978).
4. P. Dehal et al., *Science* **298**, 2157 (2002).
5. J. P. Vinson et al., *Genome Res.* **15**, 1127 (2005).
6. R. T. Hinegardner, *Anal. Biochem.* **39**, 197 (1971).
7. C. G. Elsik, A. J. Mackey, J. T. Reese, N. V. Milshina, D. S. Roos, G. M. Weinstock, *Genome Biol.*, in press.
8. Sea Urchin Genome Project (<http://supg.caltech.edu/resources/annotation.php>).
9. Genboree ([www.genboree.org](http://www.genboree.org)).
10. M. Samanta et al., *Science* **314**, 960 (2006).
11. J. C. Sullivan et al., *Nucleic Acids Res.* **34**, D495 (2006).
12. S. C. Materna, M. Howard-Ashby, R. F. Gray, E. H. Davidson, *Dev. Biol.* **10.1016/j.ydbio.2006.08.032**, in press.
13. F. Raible et al., *Dev. Biol.*, in press.
14. K. Grobner, *Verh. Zool. Bot. Ges. Wien* **58**, 491 (1908).
15. E. H. Davidson, in *Gene Regulatory Networks in Development and Evolution* (Academic Press/Elsevier, San Diego, CA, 2006).
16. E. H. Davidson et al., *Science* **295**, 1669 (2002).
17. E. H. Davidson et al., *Dev. Biol.* **246**, 162 (2002).
18. Z. Wei, R. C. Angerer, L. M. Angerer, *Dev. Biol.* **10.1016/j.ydbio.2006.08.034**, in press.
19. D. H. Erwin, E. H. Davidson, *Development* **129**, 3021 (2002).
20. M. Howard-Ashby, C. T. Brown, S. C. Materna, L. Chen, E. H. Davidson, *Dev. Biol.*, in press.
21. G. Manning, G. D. Plowman, T. Hunter, S. Sudarsanam, *Trends Biochem. Sci.* **27**, 514 (2002).
22. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, *Science* **298**, 1912 (2002).
23. C. Bradham et al., *Dev. Biol.* **10.1016/j.ydbio.2006.08.074**, in press.
24. C. Byrum et al., *Dev. Biol.*, in press.
25. P. Dehal, J. L. Boore, *PLoS Biol.* **3**, e314 (2005).
26. X. Gu, Y. Wang, J. Gu, *Nat. Genet.* **31**, 205 (2002).
27. A. McLysaght, K. Hokamp, K. H. Wolfe, *Nat. Genet.* **31**, 200 (2002).
28. W. Beane, E. Voronina, G. M. Wessel, D. R. McClay, *Dev. Biol.* **10.1016/j.ydbio.2006.08.046**, in press.
29. A. Kusserow et al., *Nature* **433**, 156 (2005).
30. J. Croce et al., *Dev. Biol.*, in press.
31. The Wnt homepage ([www.stanford.edu/~rnusse/wntwindow.html](http://www.stanford.edu/~rnusse/wntwindow.html)).
32. C. A. Whittaker et al., *Dev. Biol.* **10.1016/j.ydbio.2006.07.044**, in press.
33. Z. Pancer, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13156 (2000).
34. Z. Pancer, J. P. Rast, E. H. Davidson, *Immunogenetics* **49**, 773 (1999).
35. S. Akira, S. Uematsu, O. Takeuchi, *Cell* **124**, 783 (2006).
36. S. D. Fugmann, C. Messier, L. A. Novack, R. A. Cameron, J. P. Rast, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3728 (2006).
37. V. V. Kapitonov, J. Jurka, *PLoS Biol.* **3**, e181 (2005).
38. J. P. Rast et al., *Science* **314**, 952 (2006).
39. L. C. Smith, K. Azumi, M. Nonaka, *Immunopharmacology* **42**, 107 (1999).

40. L. C. Smith, L. A. Clow, D. P. Terwilliger, *Immunol. Rev.* **180**, 16 (2001).
41. T. Hibino *et al.*, *Dev. Biol.* 10.1016/j.ydbio.2006.08.065, in press.
42. J. Goldstone *et al.*, *Dev. Biol.* 10.1016/j.ydbio.2006.08.066, in press.
43. H. Qin, J. A. Powell-Coffman, *Dev. Biol.* **270**, 64 (2004).
44. J. A. Walisser, E. Glover, K. Pande, A. L. Liss, C. A. Bradford, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17858 (2005).
45. D. M. Duncan, E. A. Burgess, I. Duncan, *Genes Dev.* **12**, 1290 (1998).
46. W. Xie, R. M. Evans, *J. Biol. Chem.* **276**, 37739 (2001).
47. R. D. Burke *et al.*, *Dev. Biol.* 10.1016/j.ydbio.2006.08.007, in press.
48. J. L. S. Cobb, in *Nervous Systems of Invertebrates*, M. A. Ali, Ed. (Plenum, New York, 1987), pp. 483–525.
49. J. L. Cobb, V. W. Pantreath, *Tissue Cell* **9**, 125 (1977).
50. B. T. Livingston *et al.*, *Dev. Biol.* 10.1016/j.ydbio.2006.07.047, in press.
51. K. Kawasaki, T. Suzuki, K. M. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11356 (2004).
52. R. L. Morris *et al.*, *Dev. Biol.* 10.1016/j.ydbio.2006.08.052, in press.
53. E. V. Koonin, L. Aravind, *Cell Death Differ.* **9**, 394 (2002).
54. We gratefully acknowledge the following support: BCM-HGSC, National Human Genome Research Institute (NIH) grant 5 U54 HG003273; Naples Workshop, Stazione Zoologica Naples and the Network of Excellence "Marine Genomics Europe" (GOCE-04-505403); M. Elphick, Biotechnology and Biological Sciences Research Council (BBSRC), UK, grant S19916; J. Rast laboratory, Natural Sciences and Engineering Research Council (NSERC) of Canada, Canadian Institutes of Health Research (CIHR), and the Uehara Memorial Foundation; J. A. Coffman, **Mount Desert Island Biological Laboratory** (MDIBL), NIH grant GM070840; M. C. Thorndyke, K. H. Wilson, F. Hallböök, R. P. Olinski, Swedish Science Research Council, Network of Excellence Marine Genomics Europe (GOCE-04-505403), European Union Research Training Networks FP5 Trophic Neurogenome HPRN-ct-2002-00263, and the Royal Swedish Academy of Sciences, STINT; E. H. Davidson, R. A. Cameron, the Center for Computational Regulatory Genomics (E. H. Davidson, principal investigator) was supported by the NIH grant RR-15044, NSF IOB-0212869, and the Beckman Institute; also, support for the E. H. Davidson laboratory is from NIH grants HD-37105 and GM61005 and U.S. Department of Energy (DOE) grant DE-FG02-03ER63584; P. Oliveri, Camilla Chandler Frost Fellowship; G. M. Wessel laboratory supported by NSF IOB-0620607 and NIH grant R01 HD028152; B. Brandhorst, K. Bergeron, and N. Chen, NSERC; K. R. Foltz, NSF, IBN-0415581; M. Hahn, NIH grant R01ES006272; D. Burgess, NIH grant GM058231; L. C. Smith, NSF (MCB-0424235); R. O. Hynes, Howard Hughes Medical Institute and National Cancer Institute (NCI) (MIT Cancer Center core grant P30-CA14051); D. McClay, NIH grants GM61464, HD039948, and HD14483; V. D. Vacquier (group leader), G. W. Moy, H. J. Gunaratne, M. Kinukawa, M. Nomura, A. T. Neill, and Y.-H. Su, NIH grant R37-HD12896; R. D. Burke, NSERC and CIHR; L. M. Angerer, National Institute of Dental and Craniofacial Research (NIDCR), R. C. Angerer (NIDCR), Z. Wei (NIDCR), G. Humphrey, National Institute of Child Health and Human Development (NICHD), M. Landrum, National Center for Biotechnology Information (NCBI), O. Ermolaeva (NCBI), P. Kitts (NCBI), K. Pruitt (NCBI), V. Sapojnikov (NCBI), A. Souvorov (NCBI), W. Hiavina (NCBI), S. Fugmann, National Institute on Aging (NIA), M. Dean, National Cancer Institute—Frederick (NCIFCRF) Intramural Research Program of the NIH; P. Cormier, Association pour la Recherche contre le Cancer (ARC), France, grants 4247 and 3507 to P.C., Ligue Nationale contre le Cancer to P.C., Conseil Régional de Bretagne and Conseil Général du Finistère; W. H. Klein, National Eye Institute, NIH grant EY11930, NICHD HD66219, and the Robert A.

Welch Foundation (G-0010); N. Adams, NSF grant IBN 0417003 and the Department of the Navy, Office of Naval Research, under Award N00014-05-1-0855; D. Epel, NSF 0417225; A. Hamdoun, F32-HD47136; C. Byrum, American Heart Association grant 0420074Z; K. Walton, U.S. Army Medical Research and Materiel Command grant W81XWH-04-1-0324; J. Stegeman, NIH 2P42 ES07381; and J. Goldstone, NIH F32 ES012794.

#### Sea Urchin Genome Sequencing Consortium

**Overall project leadership:** Erica Sodergren,<sup>1,2</sup> George M. Weinstock,<sup>1,2</sup> Eric H. Davidson,<sup>3</sup> R. Andrew Cameron<sup>3</sup>

**Principal investigators:** Richard A. Gibbs,<sup>1,2</sup> George M. Weinstock<sup>1,2</sup>

**Annotation section leaders:** Robert C. Angerer,<sup>4</sup> Lynne M. Angerer,<sup>4</sup> Maria Ina Arnone,<sup>5</sup> David R. Burgess,<sup>6</sup> Robert D. Burke,<sup>7</sup> R. Andrew Cameron,<sup>3</sup> James A. Coffman,<sup>8</sup> Eric H. Davidson,<sup>3</sup> Michael Dean,<sup>9</sup> Maurice R. Elphick,<sup>10</sup> Charles A. Ettensohn,<sup>11</sup> Kathy R. Foltz,<sup>12</sup> Amro Hamdoun,<sup>13</sup> Richard O. Hynes,<sup>14</sup> William H. Klein,<sup>15</sup> William Marzluff,<sup>16</sup> David R. McClay,<sup>17</sup> Robert L. Morris,<sup>18</sup> Arcady Mushegian,<sup>19,20</sup> Jonathan P. Rast,<sup>21</sup> Erica Sodergren,<sup>1,2</sup> L. Courtney Smith,<sup>23</sup> Michael C. Thorndyke,<sup>24</sup> Victor D. Vacquier,<sup>24</sup> George M. Weinstock,<sup>1,2</sup> Gary M. Wessel,<sup>26</sup> Greg Wray,<sup>27</sup> Lan Zhang<sup>1,2</sup>

**Annotation: Gene list:** Erica Sodergren<sup>1,2</sup> (leader), George M. Weinstock<sup>1,2</sup> (leader), Robert C. Angerer,<sup>4</sup> Lynne M. Angerer,<sup>4</sup> R. Andrew Cameron,<sup>3</sup> Eric H. Davidson,<sup>3</sup> Christine G. Elisk,<sup>27</sup> Olga Ermolaeva,<sup>29</sup> Wratko Hlavina,<sup>29</sup> Gretchen Hofmann,<sup>30</sup> Paul Kitts,<sup>28</sup> Melissa J. Landrum,<sup>28</sup> Aaron J. Mackey,<sup>32\*</sup> Donna Maglott,<sup>28</sup> Georgia Panopoulou,<sup>33</sup> Albert J. Poustka,<sup>33</sup> Kim Pruitt,<sup>28</sup> Victor Sapojnikov,<sup>29</sup> Xingzhi Song,<sup>1,2</sup> Alexandre Souvorov,<sup>28</sup> Victor Solovjev,<sup>34</sup> Zheng Wei,<sup>4</sup> Charles A. Whittaker,<sup>35</sup> Kim Worley,<sup>1,2</sup> Lan Zhang<sup>1,2</sup>

**Assembly of genome:** Erica Sodergren<sup>1,2</sup> (leader), George M. Weinstock<sup>1,2</sup> (leader), K. James Durbin,<sup>1,2</sup> Richard A. Gibbs,<sup>1,2</sup> Yufeng Shen<sup>1,2</sup> (v 2.1), Xingzhi Song<sup>1,2</sup> (v 0.5), Kim Worley,<sup>1,2</sup> Lan Zhang<sup>1,2</sup>

**Basal transcription apparatus proteins and polymerases chromatin proteins:** Greg Wray<sup>27</sup> (leader), Olivier Fedrigo,<sup>26</sup> David Garfield,<sup>27</sup> Ralph Haygood,<sup>17</sup> Alexander Primus,<sup>26</sup> Rahul Satija,<sup>26</sup> Tonya Severson<sup>27</sup>

**BCM-HGSC annotation database and Genboree:** Lan Zhang<sup>1,2</sup> (leader), Erica Sodergren<sup>1,2</sup> (leader), George M. Weinstock<sup>1,2</sup> (leader), Manuel L. Gonzalez-Garay,<sup>1,2</sup> Andrew R. Jackson,<sup>1,2</sup> Aleksandar Milosavljevic,<sup>1,2</sup> Xingzhi Song,<sup>1,2</sup> Mark Tong,<sup>1,2</sup> Kim Worley<sup>1,2</sup>

**Biomaterialization:** Charles A. Ettensohn<sup>11</sup> (leader), R. Andrew Cameron,<sup>3</sup> Christopher E. Killian,<sup>36</sup> Melissa J. Landrum,<sup>31</sup> Brian T. Livingston,<sup>37</sup> Fred H. Wilt<sup>36</sup>

**Cell physiology:** James A. Coffman<sup>8</sup> (leader), William Marzluff<sup>16</sup> (leader), Arcady Mushegian,<sup>19,20</sup> (leader), Nikki Adams,<sup>37</sup> Robert Belle,<sup>38,39</sup> Seth Carbonneau,<sup>8</sup> Rocky Cheung,<sup>36</sup> Patrick Cormier,<sup>38,39</sup> Bertrand Cosson,<sup>38,39</sup> Jennifer Croce,<sup>1,2</sup> Antonio Fernandez-Guerra,<sup>40,41</sup> Anne-Marie Genevieve,<sup>40,41</sup> Manisha Goel,<sup>19</sup> Hemant Kelkar,<sup>42</sup> Julia Morales,<sup>38,39</sup> Odile Mulner-Lorillon,<sup>39,40</sup> Anthony J. Robertson<sup>8</sup>

**Cellular defense:** Amro Hamdoun<sup>13</sup> (leader), Jared V. Goldstone<sup>42</sup> (leader), Nikki Adams,<sup>36</sup> Bryan Cole,<sup>13</sup> Michael Dean,<sup>9</sup> David Epel,<sup>13</sup> Bert Gold,<sup>9</sup> Mark E. Hahn,<sup>43</sup> Meredith Howard-Ashby,<sup>3</sup> Mark Scally,<sup>47</sup> John J. Stegeman<sup>43</sup>

**Ciliogenesis and ciliary compounds:** Robert L. Morris<sup>18</sup> (leader), Erin L. Allgood,<sup>18</sup> Jonah Cool,<sup>18</sup> Kyle M. Judkins,<sup>18</sup> Shawn S. McCafferty,<sup>18</sup> Ashlan M. Musante,<sup>18</sup> Robert A. Obar,<sup>44†</sup> Amanda P. Rawson,<sup>18</sup> Blair J. Rossetti<sup>18</sup>

**Cytoskeletal and organelle genes:** David R. Burgess<sup>6</sup> (leader), Erin L. Allgood,<sup>18</sup> Jonah Cool,<sup>18</sup> Ian R. Gibbons,<sup>45</sup> Matthew P. Hoffman,<sup>6</sup> Kyle M. Judkins,<sup>18</sup> Andrew Leone,<sup>6</sup> Shawn S. McCafferty,<sup>18</sup> Robert L. Morris,<sup>18</sup> Ashlan M. Musante,<sup>18</sup> Robert A. Obar,<sup>44†</sup> Amanda P. Rawson,<sup>18</sup> Blair J. Rossetti,<sup>18</sup> Gary M. Wessel<sup>26</sup>

**Embryonic transcriptome:** Eric H. Davidson<sup>3</sup> (leader), R. Andrew Cameron,<sup>3</sup> Sorin Istrail,<sup>46</sup> Stefan C. Materna,<sup>3</sup> Manoj P. Samanta,<sup>47,48</sup> Viktor Stolz,<sup>47</sup> Waraporn Tongprasit,<sup>47</sup> Qiang Tu<sup>3</sup>

**Embryonic temporal expression pattern list:** Robert C. Angerer<sup>4</sup> (leader), Lynne M. Angerer<sup>4</sup> (leader), Zheng Wei<sup>4</sup>

**Echinoderm adhesion:** Richard O. Hynes<sup>14</sup> (leader), Karl-Frederik Bergeron,<sup>49</sup> Bruce P. Brandhorst,<sup>50</sup> Robert D. Burke,<sup>7</sup> Charles A. Whittaker,<sup>35</sup> James Whittle<sup>51</sup>

**Echinoderm evolution:** R. Andrew Cameron<sup>3</sup> (leader), Kevin Berney,<sup>3</sup> David J. Bottjer,<sup>51</sup> Cristina Calestani,<sup>53</sup> Eric H. Davidson,<sup>3</sup> Kevin Peterson,<sup>54</sup> Elly Chow,<sup>55</sup> Qiu Autumn Yuan<sup>55</sup>

**Genome analysis [GC content]:** Eran Elhaik,<sup>56</sup> Christine G. Elisk,<sup>28</sup> Dan Graur,<sup>56</sup> Justin T. Reese<sup>28</sup>

**Genome FPC map:** Ian Bosdet,<sup>57</sup> Shin Heesun,<sup>57</sup> Marco A. Marra,<sup>57</sup> Jacqueline Schein<sup>57</sup>

**Human genetic disease orthologs:** Michael Dean<sup>9</sup> (leader), Amro Hamdoun<sup>13</sup> (leader), The Sea Urchin Genome Sequencing Consortium

**Immunity:** Jonathan P. Rast<sup>21</sup> (leader), L. Courtney Smith<sup>23</sup> (leader), Michele K. Anderson,<sup>22</sup> Kevin Berney,<sup>3</sup> Virginia Brockton,<sup>23</sup> Katherine M. Buckley,<sup>23</sup> R. Andrew Cameron,<sup>3</sup> Avis H. Cohen,<sup>58</sup> Sebastian D. Fugmann,<sup>59</sup> Taku Hibino,<sup>21</sup> Mariano Loza-Coll,<sup>21</sup> Audrey J. Majeske,<sup>23</sup> Cynthia Messier,<sup>21</sup> Sham V. Nair,<sup>60</sup> Zeev Pancer,<sup>61</sup> David P. Terwilliger<sup>22</sup>

**Neurobiology and sensory systems:** Robert D. Burke<sup>7</sup> (leader), Maurice R. Elphick<sup>10</sup> (leader), William H. Klein<sup>15</sup> (leader), Michael C. Thorndyke<sup>24</sup> (leader), Cavit Agca,<sup>62</sup> Lynne M. Angerer,<sup>4</sup> Enrique Arboleda,<sup>5</sup> Maria Ina Arnone,<sup>5</sup> Bruce P. Brandhorst,<sup>50</sup> Nansheng Chen,<sup>50</sup> Allison M. Churcher,<sup>63</sup> F. Hallböök,<sup>64</sup> Glen W. Humphrey,<sup>65</sup> Richard O. Hynes,<sup>14</sup> Mohammed M. Idris,<sup>5</sup> Takae Kiyama,<sup>15</sup> Shuguang Liang,<sup>15</sup> Dan Mellott,<sup>60</sup> Xiquan Mu,<sup>15</sup> Greg Murray,<sup>48</sup> Robert P. Olinski,<sup>64</sup> Florian Raible,<sup>66,67</sup> Matthew Rowe,<sup>10</sup> John S. Taylor,<sup>63</sup> Kristin Tessmar-Raible,<sup>66</sup> D. Wang,<sup>63</sup> Karen H. Wilson,<sup>24</sup> Shunsuke Yaguchi<sup>7</sup>

**Reproduction:** Kathy R. Foltz<sup>12</sup> (leader), Victor D. Vacquier<sup>25</sup> (leader), Gary M. Wessel<sup>26</sup> (leader), Terry Gaasterland,<sup>25</sup> Blanca E. Galindo,<sup>67</sup> Herath J. Gunaratne,<sup>25</sup> Meredith Howard-Ashby,<sup>3</sup> Glen W. Humphrey,<sup>65</sup> Celina Juliano,<sup>26</sup> Masashi Kinukawa,<sup>25</sup> Gary W. Moy,<sup>25</sup> Anna T. Neill,<sup>25</sup> Mamoru Nomura,<sup>25</sup> Michael Raitsch,<sup>12</sup> Anna Reade,<sup>12</sup> Michelle M. Roux,<sup>12</sup> Jia L. Song,<sup>25</sup> Yi-Hsien Su,<sup>3</sup> Ian K. Townley,<sup>12</sup> Ekaterina Voronina,<sup>26</sup> Julian L. Wong<sup>26</sup>

**Sea Urchin Genome Annotation Workshop in Naples:** Maria Ina Arnone<sup>5</sup> (leader), Michael C. Thorndyke<sup>24</sup> (leader), Gabriele Amore,<sup>5</sup> Lynne M. Angerer,<sup>4</sup> Enrique Arboleda,<sup>5</sup> Margherita Branno,<sup>5</sup> Euan R. Brown,<sup>5</sup> Vincenzo Cavaliere,<sup>69</sup> Veronique Duboc,<sup>70</sup> Louise Duloquin,<sup>70</sup> Maurice R. Elphick,<sup>10</sup> Constantin Flytzanis,<sup>70,71</sup> Christian Gache,<sup>70</sup> Anne-Marie Genevieve,<sup>40,41</sup> Mohammed M. Idris,<sup>5</sup> François Lapraz,<sup>70</sup> Thierry Lepage,<sup>70</sup> Annamaria Locascio,<sup>5</sup> Pedro Martinez,<sup>73,74</sup> Giorgio Matassi,<sup>75</sup> Valeria Matranga,<sup>76</sup> David R. McClay,<sup>17</sup> Julia Morales,<sup>38,39</sup> Albert J. Poustka,<sup>33</sup> Florian Raible,<sup>66,67</sup> Ryan Range,<sup>70</sup> Francesca Rizzo,<sup>5</sup> Eric Röttinger,<sup>70</sup> Matthew Rowe,<sup>10</sup> Kristin Tessmar-Raible,<sup>66</sup> Erica Sodergren,<sup>1,2</sup> George M. Weinstock,<sup>1,2</sup> Karen Wilson<sup>24</sup>

**Signal transduction:** David R. McClay<sup>17</sup> (leader), Lynne M. Angerer,<sup>4</sup> Maria Ina Arnone,<sup>5</sup> Wendy Beane,<sup>17</sup> Cynthia Bradham,<sup>17</sup> Christine Byrum,<sup>17,78</sup> Jennifer Croce,<sup>1,2</sup> Veronique Duboc,<sup>70</sup> Louise Duloquin,<sup>70</sup> Christian Gache,<sup>70</sup> Anne-Marie Genevieve,<sup>40,41</sup> Tom Glenn,<sup>17</sup> Taku Hibino,<sup>22</sup> Sofia Hussain,<sup>37</sup> François Lapraz,<sup>70</sup> Thierry Lepage,<sup>70</sup> Brian T. Livingston,<sup>37</sup> Mariano Loza,<sup>21</sup> Gerard Manning,<sup>76</sup> Esther Miranda,<sup>17</sup> Ryan Range,<sup>70</sup> Francesca Rizzo,<sup>5</sup> Eric Röttinger,<sup>70</sup> Rebecca Thomson,<sup>17,78</sup> Katherine Walton,<sup>17</sup> Zheng Wei,<sup>4</sup> Gary M. Wessel,<sup>26</sup> Athula Wikramanayake,<sup>77</sup> Karen H. Wilson,<sup>23</sup> Charles Whittaker,<sup>35</sup> Shu-Yu Wu,<sup>17</sup> Ronghui Xu<sup>78</sup>

**Transcription regulatory factors:** Eric H. Davidson<sup>3</sup> (leader), Maria Ina Arnone,<sup>5</sup> Margherita Branno,<sup>5</sup> C. Titus Brown,<sup>3</sup> R. Andrew Cameron,<sup>3</sup> Lili Chen,<sup>3</sup> Rachel F. Gray,<sup>3</sup> Meredith Howard-Ashby,<sup>3</sup> Sorin Istrail,<sup>46</sup> Pei Yun Lee,<sup>3</sup> Annamaria Locascio,<sup>5</sup> Pedro Martinez,<sup>73,74</sup> Stefan C. Materna,<sup>3</sup> Jongmin Nam,<sup>3</sup> Paola Oliveri,<sup>3</sup> Francesca Rizzo,<sup>5</sup> Joel Smith<sup>3</sup>

**DNA sequencing:** Donna Muzny<sup>1,2</sup> (leader), Erica Sodergren<sup>1,2</sup> (leader), Richard A. Gibbs<sup>1,2</sup> (leader), George M. Weinstock<sup>1,2</sup> (leader), Stephanie Bell,<sup>1,2</sup> Joseph Chacko,<sup>1,2</sup> Andrew Cree, Stacey Curry,<sup>1,2</sup> Clay Davis,<sup>1,2</sup> Huyen Dinh,<sup>1,2</sup> Shannon Dugan-Rocha,<sup>1,2</sup> Jerry Fowler,<sup>1,2</sup> Rachel Gill,<sup>1,2</sup> Cerrisa Hamilton,<sup>1,2</sup> Judith Hernandez,<sup>1,2</sup> Sandra Hines,<sup>1,2</sup> Jennifer Hume,<sup>1,2</sup> LaRonda Jackson,<sup>1,2</sup> Angela Jolivet,<sup>1,2</sup> Christie Kovar,<sup>1,2</sup> Sandra Lee,<sup>1,2</sup> Lora Lewis,<sup>1,2</sup> George Miner,<sup>1,2</sup> Margaret Morgan,<sup>1,2</sup> Lynne V. Nazareth,<sup>1,2</sup> Geoffrey Okwuonu,<sup>1,2</sup> David Parker,<sup>1,2</sup> Ling-Ling Pu,<sup>1,2</sup> Yufeng Shen,<sup>1,2</sup> Rachel Thorn,<sup>1,2</sup> Rita Wright<sup>1,2</sup>

<sup>1</sup>Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.

<sup>2</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.

<sup>3</sup>Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.

<sup>4</sup>National Institute of Dental and Craniofacial Research, NIH, Bethesda, MD

20892, USA. <sup>5</sup>Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Napoli, Italy. <sup>6</sup>Department of Biology, Boston College, Chestnut Hill, MA 02467, USA. <sup>7</sup>Department of Biology, Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada, V8W 3N5. <sup>8</sup>Mount Desert Island Biological Laboratory, Salisbury Cove, ME 04672, USA. <sup>9</sup>Human Genetics Section, Laboratory of Genomic Diversity, National Cancer Institute—Frederick, Frederick, MD 21702, USA. <sup>10</sup>School of Biological and Chemical Sciences, Queen Mary, University of London, London E1 4NS, UK. <sup>11</sup>Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA, 15213, USA. <sup>12</sup>Department Molecular, Cellular and Developmental Biology and the Marine Science Institute, University of California, Santa Barbara, Santa Barbara, CA 93106–9610, USA. <sup>13</sup>Hopkins Marine Station, Stanford University, Pacific Grove, CA 93950, USA. <sup>14</sup>Howard Hughes Medical Institute, Center for Cancer Research, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. <sup>15</sup>Departments of Biochemistry and Molecular Biology, University of Texas, M. D. Anderson Cancer Center, Houston, TX, 77030, USA. <sup>16</sup>Molecular Biology and Biotechnology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. <sup>17</sup>Department of Biology, Duke University, Durham, NC 27708, USA. <sup>18</sup>Department of Biology, Wheaton College, Norton, MA 02766, USA. <sup>19</sup>Stowers Institute for Medical Research, Kansas City, MO 64110, USA. <sup>20</sup>Department of Microbiology, Kansas University Medical Center, Kansas City, KS 66160, USA. <sup>21</sup>Sunnybrook Research Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada M4N 3M5. <sup>22</sup>Department of Immunology, University of Toronto, Toronto, Ontario, Canada, M4N 3M5. <sup>23</sup>Department of Biological Sciences, George Washington University, Washington, DC 20052, USA. <sup>24</sup>Royal Swedish Academy of Sciences, Kristineberg Marine Research Station, Fiskebackskil, 450 34, Sweden. <sup>25</sup>Marine Biology, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093–0202, USA. <sup>26</sup>Department of Molecular and Cellular Biology and Biochemistry, Brown University Providence, RI 02912, USA. <sup>27</sup>Department of Biology and Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708, USA. <sup>28</sup>Department of Animal Science, Texas A&M University, College Station, TX 77843, USA. <sup>29</sup>National Center for Biotechnology Information, National Library of Medicine, NIH, Bethesda, MD 20894, USA. <sup>30</sup>Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA. <sup>31</sup>National Center for Biotechnology Information, NIH, Bethesda, MD 20892, USA. <sup>32</sup>Penn Genomics Institute, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>33</sup>Evolution and Development Group, Max-Planck Institut für Molekulare Genetik, 14195 Berlin, Germany. <sup>34</sup>Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. <sup>35</sup>Center for Cancer Research, MIT, Cambridge, MA 02139, USA. <sup>36</sup>Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720–3200, USA. <sup>37</sup>Department of Biology, University of South Florida, Tampa, FL 33618, USA. <sup>38</sup>Université Pierre et Marie Curie (Paris 6), UMR 7150, Equipe Cycle Cellulaire et Développement, Station Biologique de Roscoff, 29682 Roscoff Cedex, France. <sup>39</sup>CNRS, UMR 7150, Station Biologique de Roscoff, 29682 Roscoff Cedex, France. <sup>40</sup>CNRS, UMR7628, Banyuls-sur-Mer, F-66650, France. <sup>41</sup>Université Pierre et Marie Curie (Paris 6), UMR7628, Banyuls-sur-Mer, F-66650, France. <sup>42</sup>Center for Bioinformatics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. <sup>43</sup>Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. <sup>44</sup>Tethys Research, LLC, 2115 Union Street, Bangor, Maine 04401, USA. <sup>45</sup>Department of Molecular, Cellular, and Developmental Biology, University of California, Berkeley, Berkeley, CA 94720, USA. <sup>46</sup>Center for Computational Molecular Biology, and Computer Science Department, Brown University, Providence, RI 02912, USA. <sup>47</sup>Genome Research Facility, National Aeronautics and Space Administration, Ames Research Center, Moffet Field, CA 94035,

USA. <sup>48</sup>Systemix Institute, Cupertino, CA 95014, USA. <sup>49</sup>Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6. <sup>50</sup>Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6. <sup>51</sup>Department of Biology, Center for Cancer Research, MIT, Cambridge, MA 02139, USA. <sup>52</sup>Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089–0740, USA. <sup>53</sup>Department of Biology, University of Central Florida, Orlando, FL 32816–2368, USA. <sup>54</sup>Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA. <sup>55</sup>Center for Computational Regulatory Genomics, Beckman Institute, California Institute of Technology, Pasadena, CA 91125, USA. <sup>56</sup>Department of Biology and Biochemistry, University of Houston, Houston, TX 77204, USA. <sup>57</sup>Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, BC, Canada, V5Z 4E6. <sup>58</sup>Department of Biology and the Institute of Systems Research, University of Maryland, College Park, MD 20742, USA. <sup>59</sup>Laboratory of Cellular and Molecular Biology, National Institute on Aging, NIH, Baltimore, MD 21224, USA. <sup>60</sup>Department of Biological Sciences, Macquarie University, Sydney NSW 2109, Australia. <sup>61</sup>Center of Marine Biotechnology, UMBI, Columbus Center, Baltimore, MD 21202, USA. <sup>62</sup>Department of Cell Biology and Anatomy, Louisiana State University Health Sciences Center, New Orleans, LA 70112, USA. <sup>63</sup>Department of Biology, University of Victoria, Victoria, BC, Canada, V8W 2Y2. <sup>64</sup>Department of Neuroscience, Uppsala University, Uppsala, Sweden. <sup>65</sup>Laboratory of Cellular and Molecular Biophysics, National Institute of Child Health and Development, NIH, Bethesda, MD 20895, USA. <sup>66</sup>Developmental Unit, EMBL, 69117 Heidelberg, Germany. <sup>67</sup>Computational Unit, EMBL, 69117 Heidelberg, Germany. <sup>68</sup>Biotechnology Insti-

tute, Universidad Nacional Autónoma de México (UNAM), Cuernavaca, Morelos, México 62250. <sup>69</sup>Department of Cellular and Developmental Biology "Alberto Monroy," University of Palermo, 90146 Palermo, Italy. <sup>70</sup>Laboratoire de Biologie du Développement (UMR 7009), CNRS and Université Pierre et Marie Curie (Paris 6), Observatoire Océanologique, 06230 Villefranche-sur-Mer, France. <sup>71</sup>Department of Biology, University of Patras, Patras, Greece. <sup>72</sup>Department of Molecular and Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. <sup>73</sup>Departament de Genètica, Universitat de Barcelona, 08028–Barcelona, Spain. <sup>74</sup>Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. <sup>75</sup>Institut Jacques Monod, CNR-UMR 7592, 75005 Paris, France. <sup>76</sup>Consiglio Nazionale delle Ricerche, Istituto di Biomedicina e Immunologia Molecolare "Alberto Monroy," 90146 Palermo, Italy. <sup>77</sup>Razavi-Newman Center for Bioinformatics, Salk Institute for Biological Studies, La Jolla, CA 92186, USA. <sup>78</sup>Department of Zoology, University of Hawaii at Manoa, Honolulu, HI 96822, USA.

\*Present address: GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, PA 19426, USA.

†Present address: Massachusetts General Hospital Cancer Center, Charlestown, MA 02129, USA.

## Supporting Online Material

[www.sciencemag.org/cgi/content/full/314/5801/941/DC1](http://www.sciencemag.org/cgi/content/full/314/5801/941/DC1)

Materials and Methods

SOM Text

Figs. S1 to S6

Tables S1 to S8

Reference

8 August 2006; accepted 17 October 2006

10.1126/science.1133609

## REVIEW

# Genomic Insights into the Immune System of the Sea Urchin

Jonathan P. Rast,<sup>1,\*</sup> L. Courtney Smith,<sup>2</sup> Mariano Loza-Coll,<sup>1</sup> Taku Hibino,<sup>1</sup> Gary W. Litman<sup>3,4</sup>

Comparative analysis of the sea urchin genome has broad implications for the primitive state of deuterostome host defense and the genetic underpinnings of immunity in vertebrates. The sea urchin has an unprecedented complexity of innate immune recognition receptors relative to other animal species yet characterized. These receptor genes include a vast repertoire of 222 Toll-like receptors, a superfamily of more than 200 NACHT domain–leucine-rich repeat proteins (similar to nucleotide-binding and oligomerization domain (NOD) and NALP proteins of vertebrates), and a large family of scavenger receptor cysteine-rich proteins. More typical numbers of genes encode other immune recognition factors. Homologs of important immune and hematopoietic regulators, many of which have previously been identified only from chordates, as well as genes that are critical in adaptive immunity of jawed vertebrates, also are present. The findings serve to underscore the dynamic utilization of receptors and the complexity of immune recognition that may be basal for deuterostomes and predicts features of the ancestral bilaterian form.

Animal immune mechanisms are classified as acquired (adaptive), in which immune recognition specificity is the product of somatic diversification and selective clonal proliferation, or as innate, in which recognition specificity is germline encoded. Collectively, these systems act to protect the individual from invasive bacteria, viruses, and eukaryotic pathogens by detecting molecular signatures of infection and initiating effector responses. Innate immune mechanisms probably originated early

in animal phylogeny and are closely allied with wound healing and tissue maintenance functions. In many cases, their constituent elements are distributed throughout the cells of the organism. In bilaterally symmetrical animals (Bilateria), immune defense is carried out and tightly coordinated by a specialized set of mesoderm-derived cells that essentially are committed to this function (1–3). Overlaid onto this conserved core of developmental and immune programs are a variety of rapidly evolving recognition and

## The genomic repertoire for cell cycle control and DNA metabolism in *S. purpuratus*

Antonio Fernandez-Guerra <sup>a,1</sup>, Antoine Aze <sup>a</sup>, Julia Morales <sup>b</sup>, Odile Mulner-Lorillon <sup>b</sup>, Bertrand Cosson <sup>b</sup>, Patrick Cormier <sup>b</sup>, Cynthia Bradham <sup>c</sup>, Nikki Adams <sup>d</sup>, Anthony J. Robertson <sup>e</sup>, William F. Marzluff <sup>f</sup>, James A. Coffman <sup>e</sup>, Anne-Marie Genevière <sup>a,\*</sup>

<sup>a</sup> Observatoire Océanologique de Banyuls-Laboratoire Arago, CNRS-UMR7628/UPMC, 66650 Banyuls-sur-Mer, France

<sup>b</sup> Station Biologique de Roscoff, CNRS-UMR7150/UPMC, 29680 Roscoff, France

<sup>c</sup> Developmental, Cellular, and Molecular Biology Group, Duke University, Durham, NC, USA

<sup>d</sup> Department of Biological Sciences, California Polytechnic State University, San Luis Obispo, CA, USA

<sup>e</sup> Mount Desert Island Biological Laboratory, Salisbury Cove, ME, USA

<sup>f</sup> Program in Molecular Biology, University of North Carolina, Chapel Hill, NC, USA

Received for publication 30 May 2006; revised 6 September 2006; accepted 7 September 2006

Available online 14 September 2006

### Abstract

A search of the *Strongylocentrotus purpuratus* genome for genes associated with cell cycle control and DNA metabolism shows that the known repertoire of these genes is conserved in the sea urchin, although with fewer family members represented than in vertebrates, and with some cases of echinoderm-specific gene diversifications. For example, while homologues of the known cyclins are mostly encoded by single genes in *S. purpuratus* (unlike vertebrates, which have multiple isoforms), there are additional genes encoding novel cyclins of the B and K/L types. Almost all known cyclin-dependent kinases (CDKs) or CDK-like proteins have an orthologue in *S. purpuratus*; CDK3 is one exception, whereas CDK4 and 6 are represented by a single homologue, referred to as CDK4. While the complexity of the two families of mitotic kinases, Polo and Aurora, is close to that found in the nematode, the diversity of the NIMA-related kinases (NEK proteins) approaches that of vertebrates. Among the nine NEK proteins found in *S. purpuratus*, eight could be assigned orthologues in vertebrates, whereas the ninth is unique to sea urchins. Most known DNA replication, DNA repair and mitotic checkpoint genes are also present, as are homologues of the pRB (two) and p53 (one) tumor suppressors. Interestingly, the p21/p27 family of CDK inhibitors is represented by one homologue, whereas the INK4 and ARF families of tumor suppressors appear to be absent, suggesting that these evolved only in vertebrates. Our results suggest that, while the cell cycle control mechanisms known from other animals are generally conserved in sea urchin, parts of the machinery have diversified within the echinoderm lineage. The set of genes uncovered in this analysis of the *S. purpuratus* genome should enhance future research on cell cycle control and developmental regulation in this model.

© 2006 Elsevier Inc. All rights reserved.

**Keywords:** Sea urchin; Cell cycle; Replication; Mitosis; Checkpoint; Kinases

\* Corresponding author.

E-mail addresses: [Antonio@obs-banyuls.fr](mailto:Antonio@obs-banyuls.fr) (A. Fernandez-Guerra), [antoine.aze@obs-banyuls.fr](mailto:antoine.aze@obs-banyuls.fr) (A. Aze), [morales@sb-roscoff.fr](mailto:morales@sb-roscoff.fr) (J. Morales), [mulner@sb-roscoff.fr](mailto:mulner@sb-roscoff.fr) (O. Mulner-Lorillon), [cosson@sb-roscoff.fr](mailto:cosson@sb-roscoff.fr) (B. Cosson), [cormier@sb-roscoff.fr](mailto:cormier@sb-roscoff.fr) (P. Cormier), [cabrad@duke.edu](mailto:cabrad@duke.edu) (C. Bradham), [nadams@calpoly.edu](mailto:nadams@calpoly.edu) (N. Adams), [tony@mdibl.org](mailto:tony@mdibl.org) (A.J. Robertson), [marzluff@med.unc.edu](mailto:marzluff@med.unc.edu) (W.F. Marzluff), [jcoffman@mdibl.org](mailto:jcoffman@mdibl.org) (J.A. Coffman), [anne-marie.genevriere@obs-banyuls.fr](mailto:anne-marie.genevriere@obs-banyuls.fr) (A.-M. Genevière).

<sup>1</sup> Current location: Unitat de Limnologia-Departament de Biogeoquímica Aquàtica, Centre d'Estudis Avançat de Blanes, CSIC, 17300 Blanes, Catalunya, Spain.

### Introduction

The eukaryotic cell cycle orchestrates cell growth, the replication and segregation of genomic DNA, and cell division. While its core mechanisms are highly conserved, with the evolution of animals, the cell cycle became contingent on intercellular mitogenic signaling and thereby subjugated to the variety of ontogenetic programs found in different animal clades.

Progression through the eukaryotic cell cycle is controlled by a complex network of kinase activities. Central to this

machinery is the family of cyclin-dependent kinases (CDKs), which phosphorylate a set of regulatory factors. The activity of these serine/threonine kinases is in turn controlled at multiple levels: by synthesis of the regulatory cyclin subunits, by phosphorylation of both the catalytic and cyclin subunits, by binding to CDK inhibitors (CKIs), by subcellular localization, and by specific ubiquitylation targeting of cyclins and CKIs to proteasome-mediated proteolysis (Obaya and Sedivy, 2002). Cyclins involved in cell cycle control have been classified as G1-, S-, or M-phase regulators, based on when their activities are required in the cell cycle (Pines, 1999). Cyclin D-dependent kinases CDK4/CDK6 collaborate with cyclin E/CDK2 to promote phosphorylation of the retinoblastoma protein (pRB) and its relatives p107 and p130, relieving the inhibition of E2F family transcription factors and thus stimulating progression from G1 to S phase. The G1 phase of the cell cycle is characterized by low CDK activity and replication origins being assembled in a pre-replicative state competent to undergo DNA replication in the subsequent S phase. Assembly of pre-replicative complexes (pre-RC) is marked by the sequential binding of ORC (for “origin recognition complex”), Cdc6, Cdt1, and finally the DNA helicase MCM2–7 (for “mini chromosome maintenance”) complex to the origin of replication (reviewed in Bell, 2002; Spradling, 1999). Activation of CDK2 and of the heterodimeric kinase Cdc7/Dbf4 at the G1–S transition drives the cell into a state where replication can be initiated but new pre-RC cannot be assembled, ensuring that chromosomal DNA replicates only once per cell cycle (reviewed in Blow and Dutta, 2005; Woo and Poon, 2003). Completion of S phase triggers the activation of mitotic kinases which fully activate cyclin B/CDK1, initiating the complex program of chromosome segregation. Exit from mitosis is marked by the shutdown of CDK activity that follows the rapid degradation of cyclin B. Cyclin B, with other critical substrates for cell cycle progression, is targeted for degradation by the ubiquitin-dependent pathway.

The core genes involved in these biological processes have been found in animals, plants, and unicellular organisms such as yeast. However, some regulatory mechanisms have evolved differently. For example, multiple gene isotypes (e.g. those encoding cyclins) are present in certain lineages, the precise number varying between different organisms. Here we surveyed the genome of *Strongylocentrotus purpuratus* to obtain an insight into how echinoderms either conserve or have modified these fundamental systems.

The sea urchin has been an important model for many years in the study of basic biology. Crucial in developmental biology, it has also been essential in cell cycle studies, a key example being the identification of the first cyclin. It was in *Arbacia punctulata* that Tim Hunt first noticed the “unexpected behavior” of a protein synthesized early after fertilization that almost disappeared before cleavage and was “henceforth called cyclin” (Evans et al., 1983). Together with studies on amphibians, data obtained in marine invertebrates, i.e. the detection in starfish oocytes of a cytoplasmic activity which mediates maturation-inducing hormonal action and brings about germinal vesicle breakdown (Kishimoto and Kanatani, 1976),

the demonstration of the universality of this M-phase promoting factor (MPF) (Kishimoto et al., 1982), the sequencing of Clam cyclin A (Swenson et al., 1986), and the purification of cyclin B/CDK1 from starfish (Labbe et al., 1989) were decisive for understanding the mechanism underlying meiotic and mitotic onset (MPF reviewed in Kishimoto, 1999; Masui, 2001).

A key feature of sea urchins is that their gametes are stored as haploid eggs. Thus, fertilization directly triggers entry into mitosis rather than resumption of meiosis as in most animal eggs. After one cycle in which short G1 and G2 phases precede DNA replication and mitosis respectively, a rapid succession of S and M phases marks the cleavage stages. These early divisions rely on translation of maternally stored mRNA or post translational modifications of existing proteins stored in the eggs until the cell cycle becomes dependent on zygotic transcription and develops into a more conventional somatic cycle with gap phases.

The sea urchin occupies an important evolutionary position with respect to vertebrates because the echinoderms and their sister group hemichordates are the only other deuterostome animals beside the chordates. The sea urchin is thus more closely related to humans than are the other major invertebrate models *Drosophila* and *C. elegans*. The sequenced genome of *S. purpuratus* strengthens the sea urchin as a model organism for studying the regulatory networks that control cell proliferation, with potentially important implications for understanding their human counterparts.

In this paper, we describe the sea urchin repertoire for cell cycle control and DNA metabolism as identified in the genome sequence of *S. purpuratus* and supplement this description with currently available functional data.

## Materials and methods

### Identification of cell-cycle- and DNA-replication-related genes

To identify cell-cycle-related genes in the *S. purpuratus* genome, BLAST searches were performed against the Baylor College of Medicine (BCM) sea urchin database. Sequences that lack a cyclin domain according to Pfam domain analysis were eliminated. Computationally predicted models of other cell cycle genes were collected by searching the GLEAN3 gene set at Baylor, either by using BLASTP homologous amino acid sequences from known cell cycle control genes or by using the PFAM search engine set up on the Baylor annotation site. Domains were analyzed through the SMART (simple modular architectural research tool) web interface from <http://smart.embl-heidelberg.de/> or the ScanProsite search engine from <http://www.expasy.ch/>. Nek Alignment was done using ClustalW at <http://www.ebi.ac.uk/clustalw/>.

For phylogenetic trees, protein sequences were aligned with MAFFT (Katoh et al., 2005, 2002) using E-INS-i strategy. The conserved domains were selected and manually improved using Seaview (Galtier et al., 1996). In order to make a Maximum Likelihood analysis, Prottest (Abascal et al., 2005) was used to determine the best amino acid substitution matrix, RtREV+I+G+F was the most suitable for our data. This matrix was used to run the local version of PHYML (Guindon et al., 2005) to infer the ML tree. Trees were validated with a bootstrap procedure using 1000 replicates. Treeillustrator (Trooskens et al., 2005) was used to draw the final tree.

### Identification of DNA repair genes

The Gene Ontology database at Mouse Genome Informatics (MGI) (<http://www.informatics.jax.org/>) was searched with the term “DNA repair”. Under DNA repair ([GO:0006281]; {Biological process}{DNA metabolism}{DNA

repair}}), the subgroups base excision repair [GO:0006284] (17 genes, 20 annotations), double-strand break repair [GO:0006302] (8 genes, 10 annotations), nucleotide excision repair [GO:0006289] (19 genes, 21 annotations) and mismatch repair [GO:0006298] (7 genes, 8 annotations) were followed to retrieve all of the individual MGI entries. The peptide sequence for each retrieved entry was used to query the *S. purpuratus* GLEAN3 peptide sequence predictions through the BLAST server at Baylor. The results of that search are displayed in Table S2. Some mouse genes with repair annotations had no clear homolog in the GLEAN3 set and in that case the assembled *S. purpuratus* genome sequence and available ESTs were directly searched to confirm the absence. Xrcc1, not present as a base excision repair factor at MGI, was added to the base excision repair list.

## Results and discussion

### The kinase network

The cell cycle clock is operated by a network of kinase activities, at the core of which are the cyclin and CDK families. The regulatory cyclin subunits not only activate CDKs by changing the conformation at their catalytic sites, but also contribute to the selection of CDK substrates, subcellular localization, and regulation of protein stability (reviewed in Malumbres and Barbacid, 2005; Murray, 2004; Sanchez and Dynlacht, 2005; Sherr and Roberts, 2004). On top of this core machinery, additional kinases, often monomeric, regulate various events during the cell cycle. These include the mitotic kinases, aurora, polo, the NimA-related kinases (NEK), or the checkpoint kinases Chk (reviewed in Nigg, 2001). The mitogen-activated protein kinases (MAPK) signaling cascades which regulate meiotic and mitotic cycles are described in an accompanying paper (Bradham et al., 2006).

### Cyclin/CDK complexes

Cyclins were among the first characterized cell cycle regulators (Evans et al., 1983) and since their identification a large number of cyclin genes have been identified in various organisms. On the basis of sequence similarity, expression pattern, and protein activity during the cell cycle, cyclins have been grouped into several classes. In animals, at least 13 classes (A to L and T) of cyclins have been described (Pines, 1995), and the human genome encodes a total of 22 cyclins. Not all cyclins in animals are demonstrated to be involved in cell cycle regulation, and only cyclins A, B, D, and E are commonly considered to be part of the core cell cycle control machinery. A second set of cyclins, cyclins K, L and T, are involved in transcription or mRNA processing (Loyer et al., 2005), while a third set, cyclin C and H, have been implicated in regulation of both cell cycle and gene expression. The functions of the remaining cyclins, cyclin F, G, I, and J, remain somewhat unclear.

All the distinctive classes of cyclins were found in the *S. purpuratus* genome (Fig. 1), which contains at least 16 genes encoding proteins with a cyclin domain. However, in contrast to the human genome, which includes various isoforms of several of these cyclins, the sea urchin genome contains a single gene for each defined subfamily, with the exception of cyclins of the B and K/L types.

The *S. purpuratus* genome sequence confirmed that the previously identified G1–S cyclins, cyclin D and E (Moore et al., 2002; Sumerel et al., 2001), are indeed the only members of these respective subfamilies in sea urchins. Cyclin D mRNA is barely detectable in eggs and early embryos but increases dramatically at the early blastula stage, remaining thereafter at a constant level throughout embryogenesis (Moore et al., 2002). Premature expression of cyclin D in cleavage embryos is lethal; however, while cyclin D expression is required for normal development, it may not be necessary for cell cycle control. Cyclin E is synthesized during oogenesis, and the level of cyclin E protein is maintained constant throughout early sea urchin embryogenesis. This implies that fluctuations in cyclin E are not required for cell cycle progression through the early embryonic cell cycles. Instead, it has been postulated that cyclin E may play a role in the remodeling of the sperm chromatin, a prerequisite for subsequent DNA replication (Schnackenberg and Marzluff, 2002). After the blastula stage, cyclin E mRNA and protein levels are very low, and cyclin E expression is predominant only in cells that are actively dividing, suggesting that cyclin E contributes to cell cycle control at these stages.

Mitotic cyclins include three distinct gene families, cyclins A, B, and B3 (Jacobs et al., 1998), that appear to have overlapping functional properties. As in other invertebrates including starfish (Table S1), surf clam, limpet, and fly (Lehner and O'Farrell, 1989; Okano-Uchida et al., 1998), a single cyclin A has been found in sea urchin. Cyclin A is present in sea urchin oocytes, but its synthesis is not required for meiotic maturation (Voronina et al., 2003). Fertilization increases synthesis of cyclin A which associates with CDK1 and CDK2 (Moreau et al., 1998). The function of cyclin A during S phase has not yet been demonstrated in sea urchin, but with cyclin B, it drives the G2–M transition by potentiating MPF activation during early embryogenesis (Genevière A.M. et Picard A., unpublished data), as is the case in *Xenopus* (Devault et al., 1992) or starfish (Okano-Uchida et al., 2003). Moreover, cyclin A can compensate for an absence of cyclin B in early sea urchin embryos, demonstrating a functional overlap between the two mitotic cyclins (Voronina et al., 2003). Three B type cyclins were identified in *S. purpuratus*: the previously characterized sea urchin cyclin B (Lozano et al., 1998; Thatcher et al., 1995; Voronina et al., 2003), an orthologue of cyclin B3 (SPU\_15285), and a novel cyclin B related to mitotic cyclins of plants (SPU\_20986). A whole-genome analysis in *Caenorhabditis elegans* also revealed two novel B type cyclins that are closely related to the previously known cyclin B (Nieduszynski et al., 2002). A phylogenetic tree of B type cyclins (Fig. 2) shows that the novel *S. purpuratus* cyclin B is not related to any other cyclin B subfamily, a feature shared with a previously classified cyclin A-like protein in *Ciona intestinalis* (Kawashima et al., 2003), which is in fact more closely related to the sea urchin novel cyclin B than to A type cyclins (Figs. 1 and 2). These data reinforce evidence for multiple cyclin B gene duplication events (Nieduszynski et al., 2002). An early duplication event in animal evolution may have given rise to the B and B3 type cyclins, identified in all animals to date. More recent duplication events

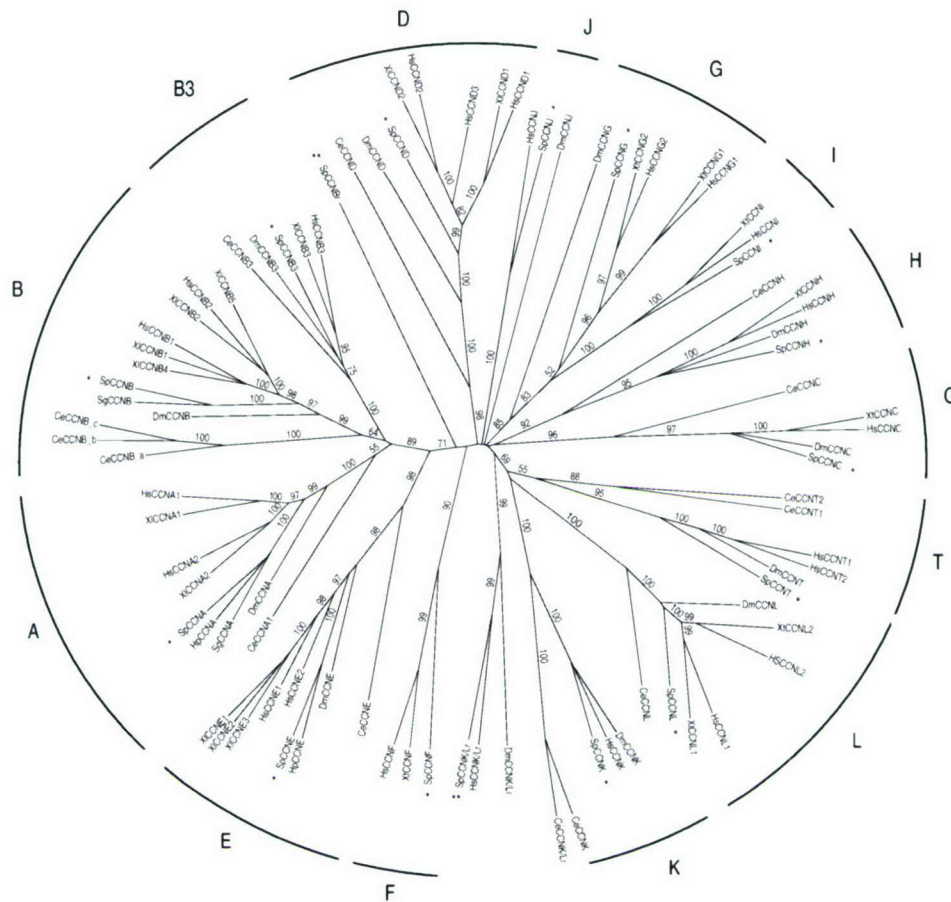


Fig. 1. Phylogenetic tree of cyclins, reconstructed by the Maximum Likelihood method (ML), with bootstrap values higher than 50% shown for each clade. Sequence Ids for the human (Hs), *Xenopus* sp. (Xl and Xt), *Drosophila melanogaster* (Dm), or *Caenorhabditis elegans* (Ce) cyclins (here noted CCN) are reported in Table S3. The *S. purpuratus* cyclins are noted with a star. The novel cyclins identified in *S. purpuratus* (Sp) which are not associated with an established cyclin subgroup are highlighted by two stars.

account for the variation in number of B type cyclins between organisms.

Cyclin C is one cyclin whose protein level does not vary significantly during the cell cycle or following growth factor stimulation. Recent analysis demonstrated that cyclin C, associated with CDK3, may play a role in the exit from quiescence, prior to pRb phosphorylation by CDK4/6/cyclin D1 in mammalian cells (Ren and Rollins, 2004). However, Cdk3 is mutated and thus non-functional in most laboratory strains of mice (Ye et al., 2001). While a potential cyclin C is present in the sea urchin, there is no apparent orthologue of the partner kinase, CDK3 (Table 1). Thus, the proposed mammalian function of cyclin C in the G0–G1 transition may not be conserved in all phyla. Cyclin C also associates with CDK8 (Leclerc and Leopold, 1996). An orthologue of CDK8 was identified in *S. purpuratus* (discussed below), which may be the only partner of cyclin C in the sea urchin. A second cyclin with a potential role both in cell cycle and transcription, cyclin H, is also present in the sea urchin genome. In vertebrates, cyclin H controls the activity of CDK7 as a Cdk activating kinase as well as its ability to phosphorylate the C-terminal domain of RNA-polymerase II (reviewed in Fisher, 2005; Harper and Elledge, 1998).

Cyclin K, T, and L constitute a protein subfamily with high sequence conservation within the conserved cyclin box and are members of the “transcription/mRNA processing” cyclin family (Loyer et al., 2005). The three cyclins have single potential orthologues in the sea urchin genome as displayed in Fig. 1. In addition, another gene of the cyclin K/L sub-type was identified in sea urchins (SPU\_023343). This gene has potential orthologues in humans (XP\_291577), mice (AAH07232), nematodes (NP\_506615), and *Drosophila* (CG31232), the encoded proteins being uncharacterized to date. This suggests that new functions of cyclin proteins remain to be discovered.

Potential orthologues of cyclin F, G, I, and J were also found in the sea urchin genome. The expression of cyclin F in sea urchin embryo, as indicated by transcriptome tiling array data (Samanta et al., 2006), is in accordance with its proposed requirement for embryonic development (Tetzlaff et al., 2004).

Since their initial discovery in yeast, the CDKs have proven to be universal regulators of the cell cycle in all eukaryotes. It is generally admitted that cell cycle progression requires a different CDK/cyclin complex for each stage of the cell cycle. Thus, as mentioned in Introduction, progression through G1 is primarily under the control of CDK4/6 family members in association with cyclin D; CDK2, paired with



Fig. 2. Phylogenetic tree for selected B type cyclins. The tree was constructed by the Maximum Likelihood method (ML), with bootstrap values shown for each clade. The *S. purpuratus* cyclin B are noted with a star. Sequence Ids for the human (Hs), *Asterina pectinifera* (Ap), *Marthasterias glacialis* (Mg), *Xenopus laevis* (Xl), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), and *Ciona intestinalis* (Ci) cyclin B (here noted CCNB) are reported in Table S3.

cyclin E and A, is required for G1–S transition and progression through S phase; while CDK1 together with cyclin A and B governs the G2–M transition. A second CDK network is critical for the regulation of transcription and mRNA processing (reviewed in Loyer et al., 2005). A few CDKs, such as CDK7, are involved in both regulatory networks. The sea urchin genome encodes most of the cell-cycle- or transcription-related CDKs encountered in mammalian cells, as displayed in Table 1. One exception is Cdk3, the absence of which was discussed above.

The core cell cycle genes Cdk1 and Cdk2 were previously identified in sea urchins (Moreau et al., 1998). Both kinases are constitutively expressed during early development. While the cyclin A and B/CDK1 activities peak in mitosis (Geneviere-Garrigues et al., 1995), CDK2 kinase activity is stable during early embryogenesis. In contrast to its role in somatic cells, CDK2 is dispensable for the activation of the first DNA replication after fertilization; however, CDK activities are required to maintain replication to one per cell cycle (Moreau et al., 1998).

Only one member of the CDK4/6 family was identified in the *S. purpuratus* genome. Similarly, a single homologue of Cdk4/6 is found in *D. melanogaster*, *C. elegans*, and *C. intestinalis*, supporting the view that the gene duplication is specific to vertebrates. Recent data from *D. melanogaster*, *C. elegans*, *S. purpuratus*, and mice suggest that cyclin D/CDK4 complexes

are not essential for cell proliferation or early development in these species and may be involved more in the regulation of cell growth and/or developmental patterning than in cell cycle progression (Datar et al., 2000; Malumbres et al., 2004; Meyer et al., 2000; Moore et al., 2002; Park and Krause, 1999).

The network of CDK/cyclin complexes implicated in transcription and mRNA processing is conserved in *S. purpuratus*. Mammalian CDK7/cyclin H and CDK8/cyclin C, which phosphorylate the C-terminal domain (CTD) of RNA-polymerase II (RNAP II), regulate in an opposite manner the initiation of transcription, belonging respectively to the TFIIF and mediator complexes (reviewed in Loyer et al., 2005). CDK9, the CTD kinase subunit of P-TEFb, associated with cyclins T or K, positively controls the elongation phase of RNAP II transcription, while CDK11 bound to cyclin L has been implicated in transcription and pre-mRNA splicing. The latest kinase in this large family to be identified, CDK12 (previously CrkRS) and its close relative CDC2L5, are the highest molecular mass CDKs identified to date (Chen et al., 2006; Ko et al., 2001; Marques et al., 2000). Recently, both proteins have been shown to regulate pre-mRNA splicing (Chen et al., 2006; Even et al., 2006). A full-length CDC2L5 cDNA was initially identified in the sea urchin *Sphaerechinus granularis* (Marques et al., 2000). In this species, the mRNA is ubiquitously expressed throughout development, and in

Table 1

*S. purpuratus* cell cycle genes identified in this study

Biochemical process/function	Family	Spu Id	Assigned name	Role of homologues	Embryonic expression
Ser/Thr kinase	CDK	002210	Sp-CDK1	G2–M	+
		007655	Sp-CDK2	G1–S	+
		Not found	Sp-CDK3		
		000151	Sp-CDK4/6	G1–S	+
		017705	Sp-CDK5	Differentiation	+
		027370	Sp-CDK7	Cell cycle/transcription	— <sup>a</sup>
		001690	Sp-CDK8	Transcription	+
		012674	Sp-CDK9	Transcription	+
		006069/026319	Sp-CDK10		+
		016748/017385	Sp-CDK11	Cell cycle/mRNA processing	+
		028236	Sp-CDK12/CDC2L5	mRNA processing	+
		016385	Sp-PCTAIRE		+
		003654	Sp-PFTAIRE		—
		003528	Sp-cyclin A	S–G2–M	+
		015285/026576	Sp-cyclin B	G2–M	+
CDK regulatory subunit	Cyclin	006444	Sp-cyclin B3	G2–M	+
		003967	Sp-cyclin C	Cell cycle/transcription	—
		007013	Sp-cyclin D	G1–S	+
		001712	Sp-cyclin E	G1–S	+
		006781/007933	Sp-cyclin F		+
		009985	Sp-cyclin G		—
		011658	Sp-cyclin H	Cell cycle/transcription	—
		007312	Sp-cyclin I		+
		021704	Sp-cyclin J		—
		005021/018199	Sp-cyclin K	Transcription	+
		014989/011295/000328	Sp-cyclin L	mRNA processing	+
		021812	Sp-cyclin T	Transcription	+
		023343	Sp-cyclin K/Lr		+
		020986	Sp-cyclin Br		+
		004789	Sp-Ccrk		+
Ser/Thr kinase	CAK	004789	Sp-Ccrk		+
Tyr kinase	Wee1	014185/027012	Sp-Wee1	S–G2–M	+
	Myt1	008280	Sp-Myt1	S–G2–M	+
Tyr phosphatase	CDC25	019568	Sp-CDC25	S–G2–M	+
Mitotic kinase	PLK	017949	Sp-PLK1	G2–M	+
		000468	Sp-PLK2	G2–M	+
		016352	Sp-PLK4	G2–M	+
		027833/027666	Sp-Aurora	G2–M	+
	Aurora NEK	019021	Sp-Nek1		+
		005277	Sp-Nek2	Mitosis	+
		023369	Sp-Nek4		+
		028348	Sp-Nek6/7		+
		005411	Sp-Nek8		+
		007456	Sp-Nek9		+
		018375/023408	Sp-Nek10		+
		014659/025176	Sp-Nek11		+
		017790	Sp-Neka		+
	Checkpoint kinase	011072/025176	Sp-ATM	DNA damage/replication	—
		011017/026783	Sp-ATR	DNA damage/replication	+
		007530	Sp-Chk1	DNA damage/replication	+
		004975	Sp-Chk2	DNA damage/replication	+
CDK/Cyclin inhibitor	Cip/Kip	005856	Sp-Ckip1	G1–S–G2–M	+
CDK/Cyclin regulator	Mat1	014250	Sp-Mat1		+
Transcription factor/cofactor	Myb	000861	Sp-Myb	G1–S	+
	Myc	003166	Sp-Myc	G1–S	+
	E2F/DP	006753	Sp-E2f3	G1–S	+
		028827	Sp-E2f4	G1–S	+
		006312	Sp-DPI	G1–S	+
		011954	Sp-Rb	G1–S	+
	Rb	004292/025615	Sp-Rbl-1	G1–S	+
		008978/008979	Sp-p53-like	G1–S	+
	Runx	006917/025612	Sp-Runt-1	G1–S	+
		007852	Sp-Runt-2	G1–S	—
	Ubiquitin-proteolysis pathway	002933	Sp-Cul-1	G1–S	+

(continued on next page)

Table 1 (continued)

Biochemical process/function	Family	Spu Id	Assigned name	Role of homologues	Embryonic expression
	Fbox	026594	Sp-Cul-2	G1-S	+
		019174/021121	Sp-Cul-3	G1-S	+
		018555/001926/009731	Sp-Cul-4	DNA damage/replication	–
		006755	Sp-Cul-5		+
		018512	Sp-Apc2	Mitosis	+
		019951	Sp-Fbw7	G1-S	–
		003152	Sp-Fbw7-like	G1-S	+
		019391	Sp-Skp2	DNA replication	+
		006781/007933	Sp-cyclin F	G2-M	+
		088809	Sp-Emi1	Mitosis	+
		0024507	Sp-Skp1	G1-S	+
		019392	Sp-Cdh1	Mitosis	+
		004395	Sp-Cdc20	Mitosis	+
		024153/008018/012580	Sp-Apc1	Mitosis	–
		022322	Sp-Apc3	Mitosis	+
		021195	Sp-Apc4	Mitosis	+
		011921	Sp-Apc5	Mitosis	+
		002193	Sp-Apc6	Mitosis	+
		009240	Sp-Apc7	Mitosis	+
		012696	Sp-Cdc23	Mitosis	+
		025744	Sp-Apc10	Mitosis	+
		021695	Sp-Apc11	Mitosis	+
		021629	Sp-SMC1	Mitosis	+
		024546	Sp-SMC2	Mitosis	+
		019803	Sp-SMC3	Mitosis	+
		013617	Sp-SMC4	Mitosis	+

<sup>a</sup> The embryonic expression (up to the mid-late gastrula stage) was deduced from the transcriptome analysis (Samanta et al., 2006). When the signal in the activity profile was negative or positive only for a minority of exons of a gene, the mRNA was considered as absent.

agreement with its proposed role in pre-mRNA splicing, the protein is localized to the nucleus (Genevière A.M., unpublished data). CDC2L5 appears to be the only gene of this type in *S. purpuratus*, suggesting that the two human kinases CDC2L5 and CrkRS represent a chordate or vertebrate-specific duplication.

While the functions of CDK/cyclin complexes regulating transitions between phases of the cell cycle or transcription are quite well understood, the binding partners and cellular role of some CDK-related proteins have still not been established. The PCTAIRE and PFTAIRE protein kinases are a subfamily of CDC2-related serine/threonine protein kinases named for a cysteine/phenylalanine-for-serine substitution in the PSTAIRE motif conserved in the initially characterized CDK proteins. Three PCTAIRE family genes have been identified in humans (Meyerson et al., 1992; Okuda et al., 1994), and two are found in mice (Okuda et al., 1992), whereas only one is present in *S. purpuratus*, as is the case in the cellular slime mold (Michaelis and Weeks, 1993). Similarly, a single PFTAIRE gene was identified in sea urchins.

#### CDK activating kinases and phosphatases

The CDK7/cyclin H complex, which phosphorylates the CTD of RNA-polymerase II, also participates in the full activation of CDKs by promoting phosphorylation of the conserved threonine residue within the T-loop region of these kinases (Kaldis, 1999). In metazoans, CDK7/cyclin H, in complex with the ring finger protein Mat1, is the major CDK activating kinase (CAK, reviewed in Harper and Elledge, 1998;

Fisher, 2005). None of these proteins has been experimentally characterized in sea urchins; however, orthologues of cyclin H, CDK7 and Mat1 are found in the *S. purpuratus* genome, suggesting that the CAK complex is functional in sea urchins.

In yeasts, a second monomeric CAK activity has been reported (Espinoza et al., 1996; Kaldis et al., 1996; Lee et al., 1999; Thuret et al., 1996). Attempts to identify orthologues of yeast CAK in the human genome were unsuccessful until the genome of *Drosophila* was used as intermediate in BLAST analysis, whereupon a protein (p42/Ccrk) with sequence similarity to both the CAK1 and CDK7 groups of CAK was discovered (Liu et al., 2004). However, its CAK activity in human remains controversial (Wohlbold et al., 2006). A homologue of p42/Ccrk is found in *S. purpuratus*. An analysis of *S. purpuratus* embryonic transcriptome (Samanta et al., 2006) indicates that this kinase is expressed during early embryogenesis, whereas neither Cdk7 nor cyclin H mRNA appears to be present. Given that Cdk7 inactivation causes embryonic lethality in *C. elegans* or *Drosophila* (Larochelle et al., 1998; Wallenfang and Seydoux, 2002), these results are intriguing. If the unexpected absence of CDK7/cyclin H mRNA in the embryo is verified, sea urchin embryogenesis may help to better understand the CAK activity requirement during embryonic cell cycles.

Cyclin B-associated CDK1 undergoes inhibitory phosphorylation at Thr 14 and Tyr 15 by the Wee1 family of protein kinases (Wee1, Mik1, and Myt1) and at entry into M phase, the inhibitory phosphates are removed by the activating phosphatase CDC25 (reviewed in O'Farrell, 2001). Structural orthologues of Wee1 and Myt1 have been identified in *S. purpuratus*, with high

similarity with the corresponding starfish *Asterina pectinifera* genes (Table S1, Nemer and Stuebing, 1996; Okano-Uchida et al., 1998; Okumura et al., 2002). Vertebrates have three CDC25 homologues (CDC25A, B and C), whereas only one was found in *S. purpuratus*, similar to its homologue in *A. pectinifera* (Deshimaru et al., 2002). Of the three vertebrate CDC25 paralogues, sea urchin CDC25 is most similar to CDC25B.

### Mitotic kinases

The process of mitosis is controlled by phosphorylation events performed not only by CDK1 but also by other serine/threonine kinases, known as mitotic kinases. These kinases include three intensely studied proteins: the polo-like kinases (PLKs), the Aurora-related kinases and the NIMA-related kinases (NEK) (Nigg, 2001).

The founding member of the PLK family, Polo, was originally identified in *D. melanogaster* where it was shown to be required for mitosis (Fenton and Glover, 1993) and to have an activity peaking cyclically at anaphase–telophase (Glover et al., 1996).

Plks have since proven to be important regulators for centrosome duplication and maturation, DNA damage checkpoint activation, mitotic onset, bipolar spindle formation, Golgi fragmentation, chromosome segregation, and cytokinesis (Barr et al., 2004; Blagden and Glover, 2003; Dai and Cogswell, 2003). Polo remains the single gene product of this family described in *Drosophila*, while four proteins sharing significant homology with Polo were identified in mammalian cells (Dai, 2005). The *S. purpuratus* genome encodes three Polo homologues (Table 1) as does the *C. elegans* genome (Chase et al., 2000; Ouyang et al., 1999), the *S. purpuratus* Plk1 being highly similar to the starfish orthologue (Table S1, Okano-Uchida et al., 2003).

The Aurora kinases also play a crucial role in regulating G2–M phase progression (Andrews et al., 2003; Eysers and Maller, 2003; Marumoto et al., 2005). The *S. purpuratus* genome apparently contains a single kinase in this family. The *C. intestinalis* genome similarly includes a single gene (Brown et al., 2004), whereas two genes have been described in *D. melanogaster* and *C. elegans*. Mammals alone have three Aurora kinases, Aurora A, B, and C, the two latter having

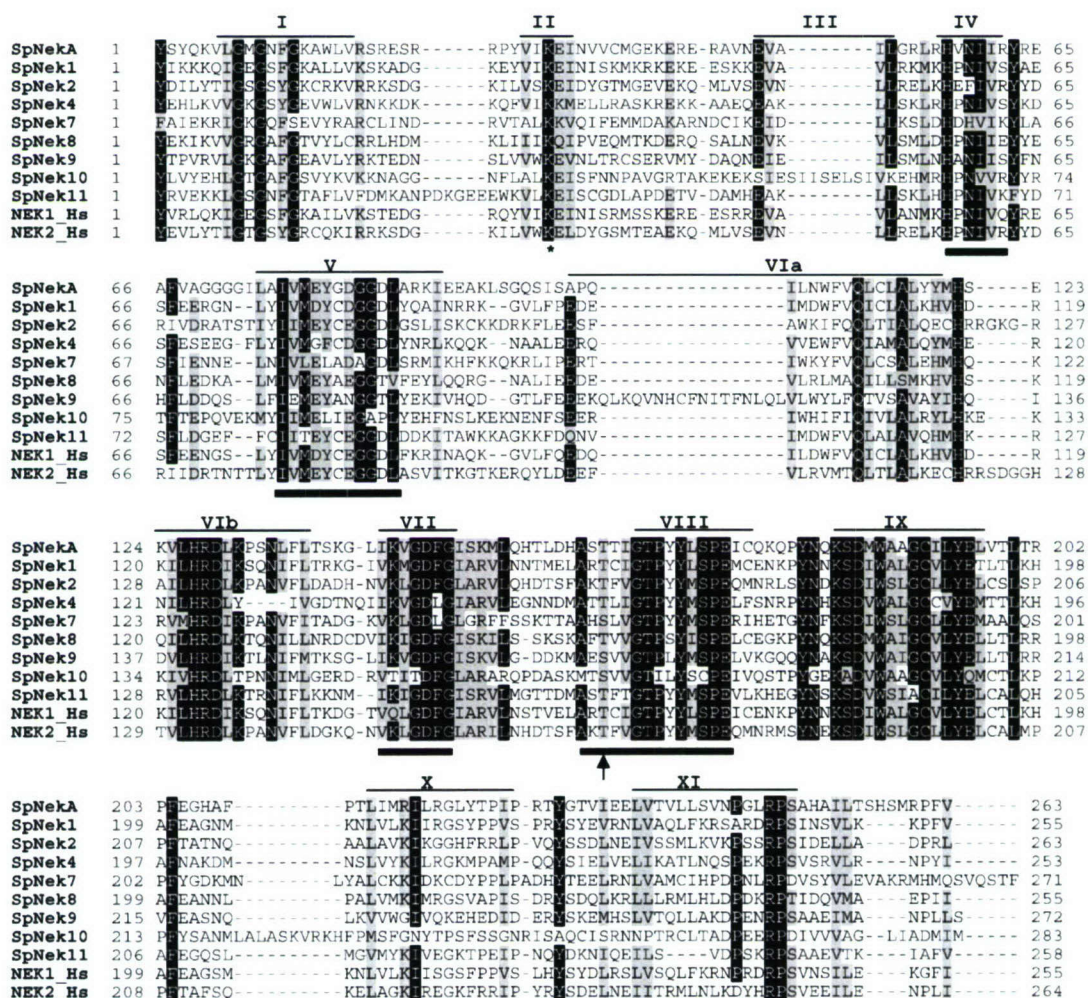


Fig. 3. Amino acid comparison of the kinase domain of the nine *S. purpuratus* NIMA-related kinases, with human Nek2 and Nek1. The eleven kinase subdomains are indicated in roman numerals, and diagnostic motifs for Nek family (Lu and Hunter, 1995) are underlined with bold bars. The star indicates the key lysine residue for kinase activity. Arrow points to a conserved serine/threonine residue, whose phosphorylation is required for activation. Identical residues are in black, conserved residues are in gray.

recently evolved from an ancestral Aurora B/C found in cold-blooded vertebrates (Brown et al., 2004). Aurora A is emerging as a critical regulator of centrosome and spindle function, while Aurora B mediates chromosome segregation by ensuring proper biorientation of sister chromatids. Aurora C, a chromosomal passenger protein similar to Aurora B, is specifically expressed in testis, indicating a role in spermatogenesis. The sequence of the sea urchin Aurora is very similar to that of its homologue in the starfish *Marthasterias glacialis* (Lapasset et al., 2005); however, like the two Auroras previously found in *C. elegans* and *D. melanogaster*, both sequences are highly divergent from the chordate counterparts (Brown et al., 2004).

The NEK proteins have been called the “third family of mitotic kinases” (O’Connell et al., 2003). The NEKs are represented in a wide variety of eukaryotic organisms, ranging from one to twenty or more representatives per genome. The NEK family possesses a highly conserved kinase domain at the N-terminus and a divergent C-terminus responsible for their specific roles in the cell. The founding member of the family,

the NIMA kinase, is required for mitotic entry and progression through mitosis in *Aspergillus nidulans*. In higher eukaryotes, NEK2, the closest relative to NIMA, plays important roles in centrosome regulation, and its overexpression leads to centriole splitting and its association with centrosome-bound proteins (Fry, 2002). The functions of other members of the NEK family are less well understood.

Nine members of the NEK family are encoded in the sea urchin genome, the kinase domains of which are aligned in Fig. 3. While eight of them clustered with the identified mammalian NEK subfamilies (Table 1, Fig. 4), no orthologue could be found for NEK3 and NEK5. The ninth sea urchin NIMA-related kinase, Sp-NEKA, displays similarity with NEK1 in its kinase domain (Fig. 4). Sp-NEKA is highly related to a sea urchin *S. granularis* cDNA (AJ841701), however, no orthologue was found in other organisms, suggesting that Sp-NEKA may constitute an echinoderm-specific innovation of the NEK family.

Quarmby and Mahjoub (2005) showed a correlation between the number of NEK kinases in the genome of an organism and

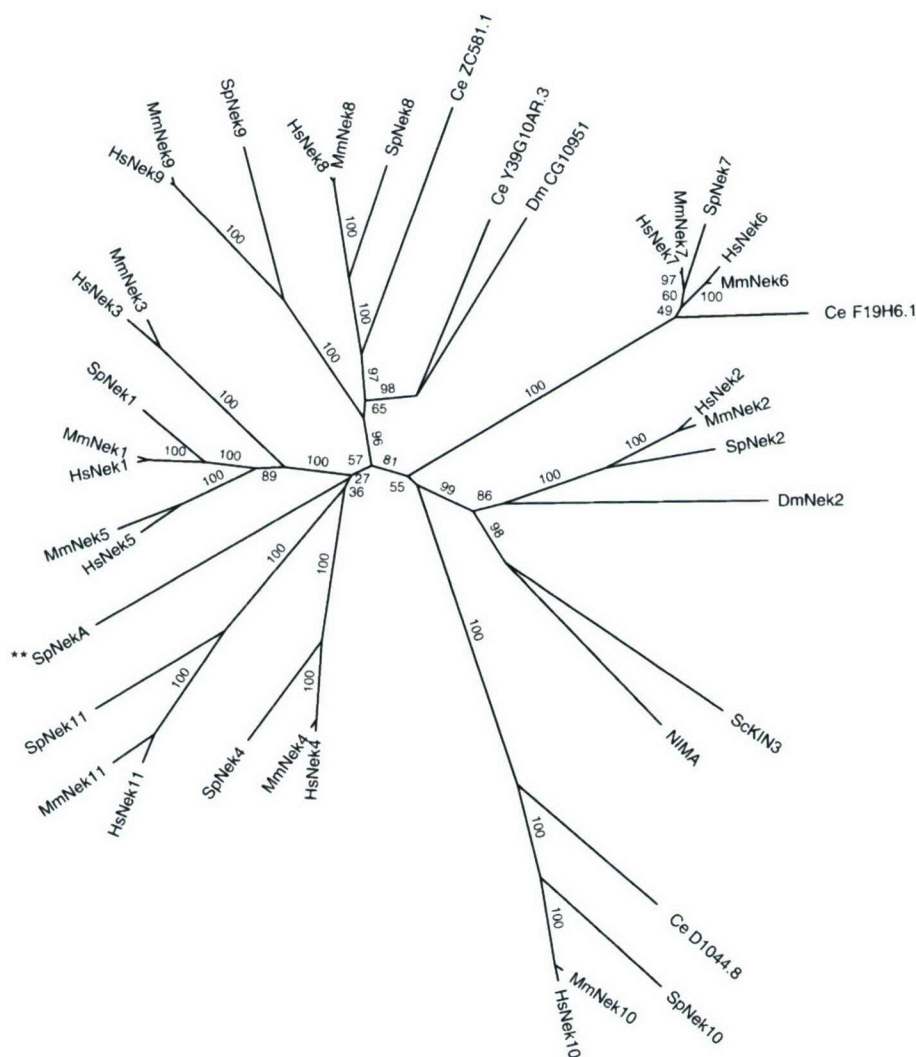


Fig. 4. Phylogenetic tree of NEK kinases. The kinase domains of NIMA-related kinases from *S. purpuratus* (Sp), human (Hs), mouse (Mm), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), and *Aspergillus nidulans* (NIMA) were identified by ScanProsite program or retrieved from the kinome dataset at Kinase.com. (\*\*) the sea urchin specific NEKA.

whether or not it has dividing ciliated cells. They proposed that the expansion of the NEK family has coevolved with centrioles, which serve both as basal bodies and microtubule-organizing centers. The complexity of NEK kinases found in *S. purpuratus* fits this hypothesis.

### Checkpoint kinases

A complex network of checkpoint pathways helps ensure precise replication and segregation of the genome by delaying cell cycle progression in response to DNA damage or replication defects, facilitating repair of the defects. Genotoxic stress arising from unreplicated DNA or from DNA damage activates DNA damage checkpoint kinases such as ATM (ataxia telangiectasia mutated) and ATR (ataxia telangiectasia and Rad3 related) (reviewed in Sancar et al., 2004). These proteins in turn activate Ser/Thr checkpoint kinases, Chk1 and Chk2, and the p53 tumor suppressor, leading to inactivation of cyclin-dependent kinases to inhibit cell progression from G1 to S (the G1–S checkpoint) or from G2 to mitosis (G2–M checkpoint) (Sancar et al., 2004; Zakian, 1995). Orthologues of the key proteins involved in both the G1–S and G2–M DNA damage checkpoints were found in *S. purpuratus* (Table 1), including one isoform each of ATM, ATR, Chk1, and Chk2. Chk1 mRNA has also been identified in *S. purpuratus* eggs by RT-PCR (Adams, unpublished data).

The CDC25 phosphatases, which trigger the final activation of the CDK/cyclin complexes, are important downstream effectors of the Chk1/Chk2 response. Depending on the cell cycle stage and the nature of the damaging agent, activated Chks phosphorylate one or the other human CDC25 isoform, promoting either ubiquitin/proteasome mediated degradation (CDC25A in G1–S or S phase checkpoint) or 14-3-3 mediated sequestration (CDC25A and C in G2–M checkpoint) of the targeted phosphatase, preventing the activation of the downstream CDK/cyclin complex (reviewed in Kristjansdottir and Rudolph, 2004; Sancar et al., 2004; Stanford and Ruderman, 2005). As mentioned above, sea urchins have a single CDC25 homologue. Three 14-3-3 genes were identified in the *S. purpuratus* genome, whereas mammals have seven (Boston and Jackson, 1980). The 14-3-3 proteins play a pivotal role in several checkpoints targeting not only the CDC25 phosphatase but also CDK2, CDK4, and the CDK inhibitor p27 at the G1–S transition, and CDK1, Wee1, Chk1, and p53 at G2–M (Hermeking and Benzinger, 2006). How specificity of interaction among 14-3-3 isoforms is generated is not entirely clear, and 14-3-3 function in the unperturbed cell cycle is still poorly understood. As the three 14-3-3 genes are expressed during early embryogenesis (as indicated by the transcriptome analysis; Samanta et al., 2006), the sea urchin embryo should provide a good model for investigating their function.

### The control of DNA replication

During S phase, the many thousands of replication forks involved in replicating chromosomal DNA must be coordinated to ensure that, despite the very large quantities of DNA

involved, chromosomal DNA is replicated once and only once per cell cycle. Cells achieve this by having distinct stages that proceed sequentially. Prior to S phase, replication origins are “licensed” by loading pre-RC onto them. The licensing of replication origins depends on the sequential recruitment to DNA of ORC, Cdc6, Cdt1 and MCM2–7 proteins (Blow and Dutta, 2005; Forsburg, 2004; Nishitani and Lygerou, 2002). The six subunits of the ORC initiator (ORC1–6) were identified in the *S. purpuratus* genome, as were the six MCM helicases. Cdt1 and Cdc6 are also present (Table 2). Recently, two additional members of the MCM helicase family were identified: MCM8 and MCM9 (Gozuacik et al., 2003; Lutzmann et al., 2005; Yoshida, 2005). While MCM8 has been proposed to function as DNA helicase with a role distinct from that of the MCM2–7 complex, MCM9 function is unknown. MCM9 is more closely related to MCM8 than to the other MCM2–7 proteins and was hitherto known only in vertebrates, and thought to be a recent duplication of the MCM8 gene (Maiorano et al., 2006). The identification of MCM9 orthologues in the genome of *S. purpuratus* and the tunicate *C. intestinalis* (Fig. 5) demonstrates that the gene is not limited to vertebrates, suggesting that the MCM8/9 duplication occurred within deuterostomes prior to the split between chordates and echinoderms.

Table 2  
*S. purpuratus* DNA replication genes identified in this study

Biochemical process/function	Family	Spu Id	Assigned name	Embryonic expression
Components of pre-replication complexes (pre-RC)	ORC	019821	Sp-ORC1	+
		010533	Sp-ORC2	+
		000544	Sp-ORC3	+
		024094	Sp-ORC4	+
		018823	Sp-ORC5	+
		005343	Sp-ORC6	+
	MCM	06096/011491	Sp-MCM2	+
		012983/006848	Sp-MCM3	+
		024515	Sp-MCM4	+
		012431	Sp-MCM5	+
		023819	Sp-MCM6	+
		026450	Sp-MCM7	+
		010595	Sp-CDC6	+
	Cdt1	002046	Sp-Cdt1	+
Inhibitor of pre-RC	Geminin	023715	Sp-Geminin	+
Components of replisome progression complexes	CDC45	023032/024816	Sp-CDC45	+
	RPA	028925	Sp-RPA1	+
		026990	Sp-RPA2	+
		019078	Sp-RPA3	+
	GINS	012376	Sp-Sld5	+
		017817/017818	Sp-Psf-1	+
		021616	Sp-Psf2	—
		018457	Sp-Psf3	+
	MCM-10	005667	Sp-MCM10	+
Polymerases (DNA-directed)	Alpha-1	000210	Sp-Pol-alpha-1	+
	Alpha2	011199	Sp-Pol-alpha-2	+
	Epsilon-1	025749	Sp-Pol-epsilon-1	+
	Epsilon-2	014691	Sp-Pol-epsilon-2	+
Regulators of initiation	CDC7	000046	Sp-CDC7	+
	Dbf4	023982	Sp-Dbf4	+
Unknown	MCM	007007	Sp-MCM8	+
		003633	Sp-MCM9	+

Activation by protein kinases is required for the initiation of licensed origins (Bell and Dutta, 2002). In higher eukaryotes, the activities of CDK2/cyclin E and Cdc7/Dbf4 have been shown to result in changes in the pre-RC that lead to the binding of Cdc45 to the MCM complex followed by the unwinding of replication origins and the recruitment of RPA, DNA polymerase  $\alpha$  and  $\epsilon$  to initiation sites. Replication initiation is tightly coupled to removal of the license of origins (reviewed in Blow and Dutta, 2005). Metazoans mainly prevent re-licensing during S phase and G2 by the inhibitory binding of geminin to Cdt1. Homologues of all the components of this DNA replication regulatory network were found in the *S. purpuratus* genome as reported in Table 2.

Additional levels of cell cycle regulation

Proliferative disorders such as cancer are associated with somatic mutations and genomic instability, which are generally caused by errors in DNA replication or mitosis. Premature entry into either S or M phase increases the probability of error, and hence multiple levels of cell cycle control machinery are dedicated to ensuring that this does not occur. In animals, these control networks are linked to the transcriptional regulatory

networks that control development (reviewed in Coffman, 2004). In addition to being under the control of the cyclins, CDKs, and other regulatory machinery discussed above, G1–S phase and G2–M phase transitions are regulated by CDK inhibitors such as mammalian p27<sup>kip1</sup> and p21<sup>cip1/waf1</sup> and the INK4 proteins and by expression levels and activities of transcription factors and cofactors such as E2F, pRB, and p53 family members (reviewed in Coffman, 2004; Sherr, 2000). Many of the CDKs discussed above function as switches in the proteolytic pathways that control levels of cyclins and other cell cycle control proteins, phosphorylated forms of which are targeted for proteolytic degradation by ubiquitin-ligases in the SCF and APC complexes (Reed, 2003). CDK inhibitors and regulators, transcription factors associated with cell cycle control, and proteins involved in ubiquitin-mediated proteolysis found encoded in the *S. purpuratus* genome are listed in Table 1. Notable absences include the INK4 CDK inhibitors, and the ARF tumor suppressor, which to date have been identified only in mammals and are probably vertebrate (and possibly mammalian) novelties. Moreover, the presence of two pRB family homologues, two E2Fs, a single p53 homologue, and one p21<sup>cip1</sup>/p27<sup>kip1</sup> homologue in sea urchins mirrors the situation in flies, suggesting that this is the primitive bilaterian

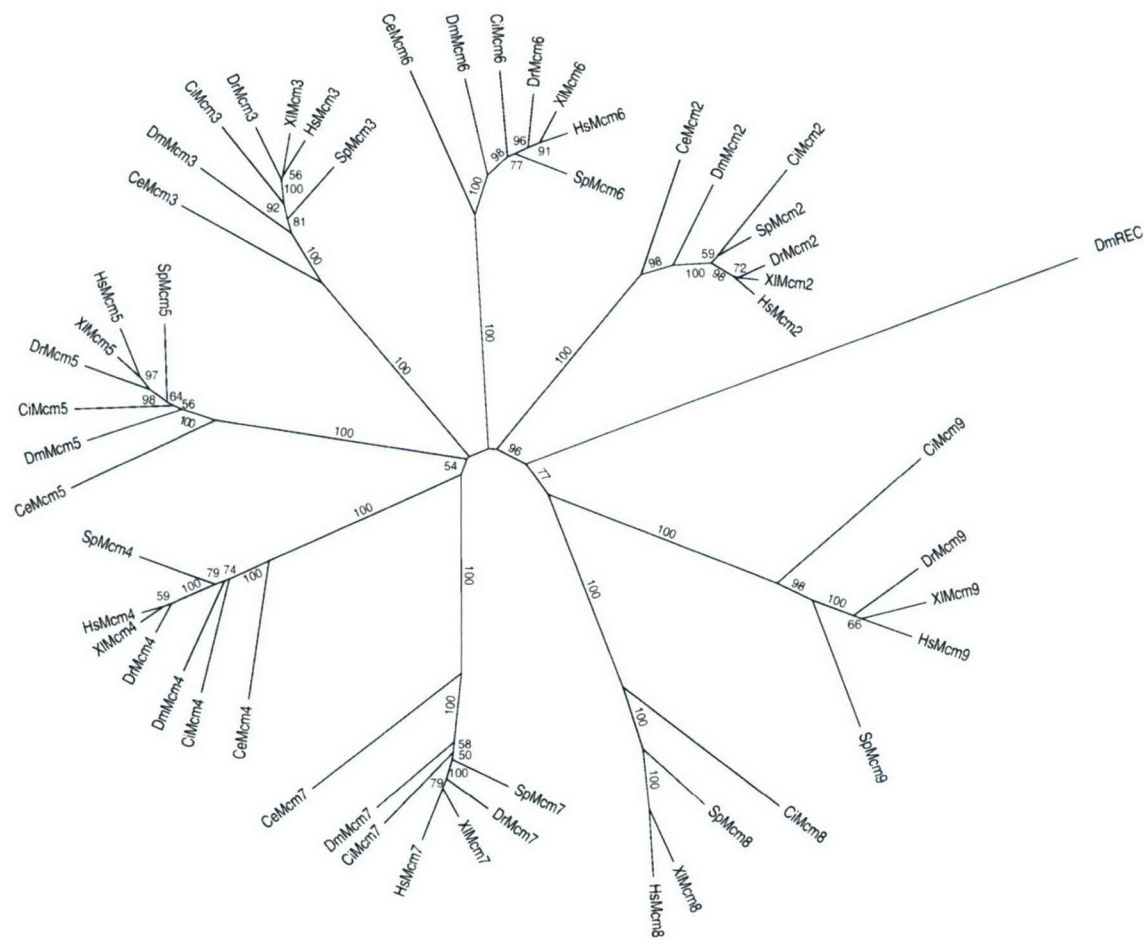


Fig. 5. Phylogenetic tree of MCM proteins. The tree was constructed by the Maximum Likelihood method (ML), with bootstrap values shown for each clade. Sequence Ids for the human (Hs), *Xenopus laevis* (XI), *Danio rerio* (Dr), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), and *Ciona intestinalis* (Ci) MCMs are reported in Table S3.

repertoire of these genes, which underwent expansion and diversification in the chordate–vertebrate lineage.

An extensive network of enzymes and other proteins that detect and repair damaged DNA provides another critical line of defense against mutation and genomic instability. Sea urchin genes associated with DNA repair functions (base excision, double-strand break, nucleotide excision, and mismatch repair) were identified using BLAST searches with mouse protein sequences retrieved by Gene Ontology searches (as described in Materials and methods); these are listed in Table S2. In general (and as might be expected), this set of genes is highly conserved between vertebrates and sea urchins. Homologues were found for all of the searched mammalian repair proteins, with the exception of one DNA base excision repair protein and two nucleotide excision repair proteins (Table S2).

## Conclusions

The *S. purpuratus* genomic toolkit for cell cycle control is similar to, albeit somewhat simpler than, that of vertebrates. For most gene families, each subtype is represented by a single gene with some limited but notable exceptions. This holds particularly true for the CDK and cyclin families.

A comparison of sea urchin cell cycle genes with those from vertebrates, *C. elegans*, *Drosophila melanogaster*, as well as the tunicate *C. intestinalis*, provides new insight into the evolution of cell cycle control. While some gene subtypes have undergone lineage-specific expansions in vertebrates (i.e. the cyclins A, B, D, and E, the mitotic kinases aurora, polo, and NEK), others seem to be lost in that lineage (i.e. the novel cyclin B identified in *S. purpuratus* and *C. intestinalis* which are not found in zebrafish, mouse, or humans). On the other hand, the recently identified replicative helicase MCM9, which was previously thought to be a vertebrate innovation, is also found in sea urchins. Moreover, analysis of the sea urchin genome highlighted a few genes which either are specific to the echinoderm lineage, such as NekaA, or were identified in *S. purpuratus* and found to have potential orthologues in other species, such as the novel family of cyclin K/L-related proteins. Most of the genes participating in the initiation and progression of DNA replication, in DNA repair, or in the tight control of cell cycle events, such as the checkpoint proteins, are highly conserved between echinoderms and vertebrates. The CDK inhibitors are a notable exception, with only one p21/p27 homologue, and an absence of the INK4 family, which is apparently confined to vertebrates.

Most of the genes included in this analysis (Tables 1 and 2) are expressed during embryogenesis (109/120) as deduced from the transcriptome analysis (Samanta et al., 2006), and those strictly related to cell cycle control are all being expressed at this stage. The genomic repertoire of cell cycle regulators uncovered in this survey will thus provide molecular tools that further enhance the physiological features that have made the early sea urchin embryo a productive model for understanding cell cycle control and its developmental regulation.

## Acknowledgments

We especially thank Yves Desdevise for very helpful discussion about phylogenetic analysis and Charles A. Whittaker for providing comments on his annotation of the 14-3-3 family of proteins. Parts of this work were supported by NIH grant GM070840 (JAC), EC grant NoE Marine genomics 505403. A.F.-G. was recipient of a European predoctoral fellowship from the contract HPMT-CT-2000-00211.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2006.09.012.

## References

- Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21, 2104–2105.
- Andrews, P.D., Knatko, E., Moore, W.J., Swedlow, J.R., 2003. Mitotic mechanics: the auroras come into view. *Curr. Opin. Cell Biol.* 15, 672–683.
- Barr, F.A., Sillje, H.H., Nigg, E.A., 2004. Polo-like kinases and the orchestration of cell division. *Nat. Rev., Mol. Cell Biol.* 5, 429–440.
- Bell, S.P., 2002. The origin recognition complex: from simple origins to complex functions. *Genes Dev.* 16, 659–672.
- Bell, S.P., Dutta, A., 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71, 333–374.
- Blagden, S.P., Glover, D.M., 2003. Polar expeditions—Provisioning the centrosome for mitosis. *Nat. Cell Biol.* 5, 505–511.
- Blow, J.J., Dutta, A., 2005. Preventing re-replication of chromosomal DNA. *Nat. Rev., Mol. Cell Biol.* 6, 476–486.
- Boston, P., Jackson, P., 1980. Purification and properties of a brain-specific protein, human 14-3-3 protein. *Biochem. Soc. Trans.* 8, 617–618.
- Bradham, C.A., Wendy, S.B., Arnone, I., Rizzo, F., J., C., Mushigian, A., Goel, M., Morales, J., Genevieve, A.M., Lapraz, F., Robertson, A.J., Kelkar, H., Loza-Coll, M., Townley, I.R., Raisch, M., Roux, M.M., Lepage, T., Gache, C., McClay, D.R., Foltz, K., 2006. The sea urchin kinome: a first look. *Dev. Biol.* (this issue). doi: 10.1016/j.ydbio.2006.08.074.
- Brown, J.R., Koretke, K.K., Birkeland, M.L., Sanseau, P., Patrick, D.R., 2004. Evolutionary relationships of Aurora kinases: implications for model organism studies and the development of anti-cancer drugs. *BMC Evol. Biol.* 4, 39.
- Chase, D., Golden, A., Heidecker, G., Ferris, D.K., 2000. *Caenorhabditis elegans* contains a third polo-like kinase gene. *DNA Seq.* 11, 327–334.
- Chen, H.H., Wang, Y.C., Fann, M.J., 2006. Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. *Mol. Cell Biol.* 26, 2736–2745.
- Coffman, J.A., 2004. Cell cycle development. *Dev. Cell.* 6, 321–327.
- Dai, W., 2005. Polo-like kinases, an introduction. *Oncogene* 24, 214–216.
- Dai, W., Cogswell, J.P., 2003. Polo-like kinases and the microtubule organization center: targets for cancer therapies. *Prog. Cell Cycle Res.* 5, 327–334.
- Datar, S.A., Jacobs, H.W., de la Cruz, A.F., Lehner, C.F., Edgar, B.A., 2000. The *Drosophila* cyclin D–Cdk4 complex promotes cellular growth. *EMBO J.* 19, 4543–4554.
- Deshimaru, S., Miyake, Y., Ohmiya, T., Tatsu, Y., Endo, Y., Yumoto, N., Toraya, T., 2002. Heterologous expression and catalytic properties of the C-terminal domain of starfish cdc25 dual-specificity phosphatase, a cell cycle regulator. *J. Biochem. (Tokyo)* 131, 705–712.
- Devault, A., Fesquet, D., Cavadore, J.C., Garrigues, A.M., Labbe, J.C., Lorca, T., Picard, A., Philippe, M., Doree, M., 1992. Cyclin A potentiates maturation-promoting factor activation in the early *Xenopus* embryo via inhibition of the tyrosine kinase that phosphorylates cdc2. *J. Cell Biol.* 118, 1109–1120.
- Espinoza, F.H., Farrell, A., Erdjument-Bromage, H., Tempst, P., Morgan, D.O.,

1996. A cyclin-dependent kinase-activating kinase (CAK) in budding yeast unrelated to vertebrate CAK. *Science* 273, 1714–1717.
- Evans, T., Rosenthal, E.T., Youngblom, J., Distel, D., Hunt, T., 1983. Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell* 33, 389–396.
- Even, Y., Durieux, S., Escande, M.L., Lozano, J.C., Peaucellier, G., Weil, D., Genevieve, A.M., 2006. CDC2L5, a Cdk-like kinase with RS domain, interacts with the ASF/SF2-associated protein p32 and affects splicing in vivo. *J. Cell. Biochem.* 99 (3), 890–904.
- Eyers, P.A., Maller, J.L., 2003. Regulating the regulators: Aurora A activation and mitosis. *Cell Cycle* 2, 287–289.
- Fenton, B., Glover, D.M., 1993. A conserved mitotic kinase active at late anaphase–telophase in syncytial *Drosophila* embryos. *Nature* 363, 637–640.
- Fisher, R.P., 2005. Secrets of a double agent: CDK7 in cell-cycle control and transcription. *J. Cell Sci.* 118, 5171–5180.
- Forsburg, S.L., 2004. Eukaryotic MCM proteins: beyond replication initiation. *Microbiol. Mol. Biol. Rev.* 68, 109–131.
- Fry, A.M., 2002. The Nek2 protein kinase: a novel regulator of centrosome structure. *Oncogene* 21, 6184–6194.
- Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO\_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* 12, 543–548.
- Genevieve-Garrigues, A.M., Barakat, A., Doree, M., Moreau, J.L., Picard, A., 1995. Active cyclin B-cdc2 kinase does not inhibit DNA replication and cannot drive prematurely fertilized sea urchin eggs into mitosis. *J. Cell Sci.* 108 (Pt. 7), 2693–2703.
- Glover, D.M., Ohkura, H., Tavares, A., 1996. Polo kinase: the choreographer of the mitotic stage? *J. Cell Biol.* 135, 1681–1684.
- Gozuacik, D., Chami, M., Lagorce, D., Faivre, J., Murakami, Y., Poch, O., Biemann, E., Knippers, R., Brechot, C., Paterlini-Brechot, P., 2003. Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Res.* 31, 570–579.
- Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online—A web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 33, W557–W559.
- Harper, J.W., Elledge, S.J., 1998. The role of Cdk7 in CAK function, a retro-retrospective. *Genes Dev.* 12, 285–289.
- Hermeking, H., Benzinger, A., 2006. 14-3-3 proteins in cell cycle regulation. *Semin. Cancer Biol.* 16 (3), 183–192.
- Jacobs, H.W., Knoblich, J.A., Lehner, C.F., 1998. *Drosophila* Cyclin B3 is required for female fertility and is dispensable for mitosis like Cyclin B. *Genes Dev.* 12, 3741–3751.
- Kaldis, P., 1999. The cdk-activating kinase (CAK): from yeast to mammals. *Cell. Mol. Life Sci.* 55, 284–296.
- Kaldis, P., Sutton, A., Solomon, M.J., 1996. The Cdk-activating kinase (CAK) from budding yeast. *Cell* 86, 553–564.
- Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33, 511–518.
- Kawashima, T., Tokutaka, M., Awazu, S., Satoh, N., Satou, Y., 2003. A genome-wide survey of developmentally relevant genes in *Ciona intestinalis*. VIII. Genes for PI3K signaling and cell cycle. *Dev. Genes Evol.* 213, 284–290.
- Kishimoto, T., 1999. Activation of MPF at meiosis reinitiation in starfish oocytes. *Dev. Biol.* 214, 1–8.
- Kishimoto, T., Kanatani, H., 1976. Cytoplasmic factor responsible for germinal vesicle breakdown and meiotic maturation in starfish oocyte. *Nature* 260, 321–322.
- Kishimoto, T., Kuriyama, R., Kondo, H., Kanatani, H., 1982. Generality of the action of various maturation-promoting factors. *Exp. Cell Res.* 137, 121–126.
- Ko, T.K., Kelly, E., Pines, J., 2001. CrkRS: a novel conserved Cdc2-related protein kinase that colocalises with SC35 speckles. *J. Cell Sci.* 114, 2591–2603.
- Kristjansdottir, K., Rudolph, J., 2004. Cdc25 phosphatases and cancer. *Chem. Biol.* 11, 1043–1051.
- Labbe, J.C., Capony, J.P., Caput, D., Cavadore, J.C., Derancourt, J., Kaghad, M., Lelias, J.M., Picard, A., Doree, M., 1989. MPF from starfish oocytes at first meiotic metaphase is a heterodimer containing one molecule of cdc2 and one molecule of cyclin B. *EMBO J.* 8, 3053–3058.
- Lapasset, L., Pradet-Balade, B., Lozano, J.C., Peaucellier, G., Picard, A., 2005. Nuclear envelope breakdown may deliver an inhibitor of protein phosphatase 1 which triggers cyclin B translation in starfish oocytes. *Dev. Biol.* 285, 200–210.
- Larochelle, S., Pandur, J., Fisher, R.P., Salz, H.K., Suter, B., 1998. Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev.* 12, 370–381.
- Leclerc, V., Leopold, P., 1996. The cyclin C/Cdk8 kinase. *Prog. Cell Cycle Res.* 2, 197–204.
- Lee, K.M., Saiz, J.E., Barton, W.A., Fisher, R.P., 1999. Cdc2 activation in fission yeast depends on Mcs6 and Csk1, two partially redundant Cdk-activating kinases (CAKs). *Curr. Biol.* 9, 441–444.
- Lehner, C.F., O'Farrell, P.H., 1989. Expression and function of *Drosophila* cyclin A during embryonic cell cycle progression. *Cell* 56, 957–968.
- Liu, Y., Wu, C., Galaktionov, K., 2004. p42, a novel cyclin-dependent kinase-activating kinase in mammalian cells. *J. Biol. Chem.* 279, 4507–4514.
- Loyer, P., Trembley, J.H., Katona, R., Kidd, V.J., Lahti, J.M., 2005. Role of CDK/cyclin complexes in transcription and RNA splicing. *Cell Signalling* 17, 1033–1051.
- Lozano, J.C., Schatt, P., Marques, F., Peaucellier, G., Fort, P., Feral, J.P., Genevieve, A.M., Picard, A., 1998. A presumptive developmental role for a sea urchin cyclin B splice variant. *J. Cell Biol.* 140, 283–293.
- Lu, K.P., Hunter, T., 1995. The NIMA kinase: a mitotic regulator in *Aspergillus nidulans* and vertebrate cells. *Prog. Cell Cycle Res.* 1, 187–205.
- Lutzmann, M., Maiorano, D., Mechali, M., 2005. Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2–8 protein family. *Gene* 362, 51–56.
- Maiorano, D., Lutzmann, M., Mechali, M., 2006. MCM proteins and DNA replication. *Curr. Opin. Cell Biol.* 18, 130–136.
- Malumbres, M., Barbacid, M., 2005. Mammalian cyclin-dependent kinases. *Trends Biochem. Sci.* 30, 630–641.
- Malumbres, M., Sotillo, R., Santamaria, D., Galan, J., Cerezo, A., Ortega, S., Dubus, P., Barbacid, M., 2004. Mammalian cells cycle without the D-type cyclin-dependent kinases Cdk4 and Cdk6. *Cell* 118, 493–504.
- Marques, F., Moreau, J.L., Peaucellier, G., Lozano, J.C., Schatt, P., Picard, A., Callebaut, I., Perret, E., Genevieve, A.M., 2000. A new subfamily of high molecular mass CDC2-related kinases with PITAI/VRE motifs. *Biochem. Biophys. Res. Commun.* 279, 832–837.
- Marumoto, T., Zhang, D., Saya, H., 2005. Aurora-A—A guardian of poles. *Nat. Rev., Cancer* 5, 42–50.
- Masui, Y., 2001. From oocyte maturation to the in vitro cell cycle: the history of discoveries of Maturation-Promoting Factor (MPF) and Cytostatic Factor (CSF). *Differentiation* 69, 1–17.
- Meyer, C.A., Jacobs, H.W., Datar, S.A., Du, W., Edgar, B.A., Lehner, C.F., 2000. *Drosophila* Cdk4 is required for normal growth and is dispensable for cell cycle progression. *EMBO J.* 19, 4533–4542.
- Meyerson, M., Enders, G.H., Wu, C.L., Su, L.K., Gorka, C., Nelson, C., Harlow, E., Tsai, L.H., 1992. A family of human cdc2-related protein kinases. *EMBO J.* 11, 2909–2917.
- Michaelis, C., Weeks, G., 1993. The isolation from a unicellular organism, *Dictyostelium discoideum*, of a highly-related cdc2 gene with characteristics of the PCTAIRE subfamily. *Biochim. Biophys. Acta* 1179, 117–124.
- Moore, J.C., Sumerel, J.L., Schnackenberg, B.J., Nichols, J.A., Wikramanayake, A., Wessel, G.M., Marzluff, W.F., 2002. Cyclin D and cdk4 are required for normal development beyond the blastula stage in sea urchin embryos. *Mol. Cell. Biol.* 22, 4863–4875.
- Moreau, J.L., Marques, F., Barakat, A., Schatt, P., Lozano, J.C., Peaucellier, G., Picard, A., Genevieve, A.M., 1998. Cdk2 activity is dispensable for the onset of DNA replication during the first mitotic cycles of the sea urchin early embryo. *Dev. Biol.* 200, 182–197.

- Murray, A.W., 2004. Recycling the cell cycle: cyclins revisited. *Cell* 116, 221–234.
- Nemer, M., Stuebing, E.W., 1996. WEE1-like CDK tyrosine kinase mRNA level is regulated temporally and spatially in sea urchin embryos. *Mech. Dev.* 58, 75–88.
- Nieduszynski, C.A., Murray, J., Carrington, M., 2002. Whole-genome analysis of animal A- and B-type cyclins. *Genome Biol.* 3 (RESEARCH0070).
- Nigg, E.A., 2001. Mitotic kinases as regulators of cell division and its checkpoints. *Nat. Rev., Mol. Cell Biol.* 2, 21–32.
- Nishitani, H., Lygerou, Z., 2002. Control of DNA replication licensing in a cell cycle. *Genes Cells* 7, 523–534.
- O'Connell, M.J., Krien, M.J., Hunter, T., 2003. Never say never. The NIMA-related protein kinases in mitotic control. *Trends Cell Biol.* 13, 221–228.
- O'Farrell, P.H., 2001. Triggering the all-or-nothing switch into mitosis. *Trends Cell Biol.* 11, 512–519.
- Obaya, A.J., Sedivy, J.M., 2002. Regulation of cyclin–Cdk activity in mammalian cells. *Cell. Mol. Life Sci.* 59, 126–142.
- Okano-Uchida, T., Sekiai, T., Lee, K., Okumura, E., Tachibana, K., Kishimoto, T., 1998. In vivo regulation of cyclin A/Cdc2 and cyclin B/Cdc2 through meiotic and early cleavage cycles in starfish. *Dev. Biol.* 197, 39–53.
- Okano-Uchida, T., Okumura, E., Iwashita, M., Yoshida, H., Tachibana, K., Kishimoto, T., 2003. Distinct regulators for Plk1 activation in starfish meiotic and early embryonic cycles. *EMBO J.* 22, 5633–5642.
- Okuda, T., Cleveland, J.L., Downing, J.R., 1992. PCTAIRE-1 and PCTAIRE-3, two members of a novel cdc2/CDC28-related protein kinase gene family. *Oncogene* 7, 2249–2258.
- Okuda, T., Valentine, V.A., Shapiro, D.N., Downing, J.R., 1994. Cloning of genomic loci and chromosomal localization of the human PCTAIRE-1 and -3 protein kinase genes. *Genomics* 21, 217–221.
- Okumura, E., Fukuhara, T., Yoshida, H., Hanada, S., Kozutsumi, R., Mori, M., Tachibana, K., Kishimoto, T., 2002. Akt inhibits Myt1 in the signalling pathway that leads to meiotic G2/M-phase transition. *Nat. Cell Biol.* 4, 111–116.
- Ouyang, B., Wang, Y., Wei, D., 1999. *Caenorhabditis elegans* contains structural homologs of human prk and plk. *DNA Seq.* 10, 109–113.
- Park, M., Krause, M.W., 1999. Regulation of postembryonic G(1) cell cycle progression in *Caenorhabditis elegans* by a cyclin D/CDK-like complex. *Development* 126, 4849–4860.
- Pines, J., 1995. Cyclins and cyclin-dependent kinases: a biochemical view. *Biochem. J.* 308 (Pt. 3), 697–711.
- Pines, J., 1999. Four-dimensional control of the cell cycle. *Nat. Cell Biol.* 1, E73–E79.
- Quarumby, L.M., Mahjoub, M.R., 2005. Caught Nek-ing: cilia and centrioles. *J. Cell Sci.* 118, 5161–5169.
- Reed, S.I., 2003. Ratchets and clocks: the cell cycle, ubiquitylation and protein turnover. *Nat. Rev., Mol. Cell Biol.* 4, 855–864.
- Ren, S., Rollins, B.J., 2004. Cyclin C/cdk3 promotes Rb-dependent G0 exit. *Cell* 117, 239–251.
- Samanta, M.P., Tongprasit, W., Istrail, S., Cameron, A., Tu, Q., Davidson, E.H., Stolc, V., 2006. A high-resolution transcriptome map of the sea urchin embryo. *Science* (in press).
- Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K., Linn, S., 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73, 39–85.
- Sanchez, I., Dynlacht, B.D., 2005. New insights into cyclins, CDKs, and cell cycle control. *Semin. Cell Dev. Biol.* 16, 311–321.
- Schnackenberg, B.J., Marzluff, W.F., 2002. Novel localization and possible functions of cyclin E in early sea urchin development. *J. Cell Sci.* 115, 113–121.
- Sherr, C.J., 2000. The Pezcoller lecture: cancer cell cycles revisited. *Cancer Res.* 60, 3689–3695.
- Sherr, C.J., Roberts, J.M., 2004. Living with or without cyclins and cyclin-dependent kinases. *Genes Dev.* 18, 2699–2711.
- Spradling, A.C., 1999. ORC binding, gene amplification, and the nature of metazoan replication origins. *Genes Dev.* 13, 2619–2623.
- Stanford, J.S., Ruderman, J.V., 2005. Changes in regulatory phosphorylation of Cdc25C Ser287 and Wee1 Ser549 during normal cell cycle progression and checkpoint arrests. *Mol. Biol. Cell* 16, 5749–5760.
- Sumerel, J.L., Moore, J.C., Schnackenberg, B.J., Nichols, J.A., Canman, J.C., Wessel, G.M., Marzluff, W.F., 2001. Cyclin E and its associated cdk activity do not cycle during early embryogenesis of the sea Urchin. *Dev. Biol.* 234, 425–440.
- Swenson, K.I., Farrell, K.M., Ruderman, J.V., 1986. The clam embryo protein cyclin A induces entry into M phase and the resumption of meiosis in *Xenopus* oocytes. *Cell* 47, 861–870.
- Tetzlaff, M.T., Bai, C., Finegold, M., Wilson, J., Harper, J.W., Mahon, K.A., Elledge, S.J., 2004. Cyclin F disruption compromises placental development and affects normal cell cycle execution. *Mol. Cell Biol.* 24, 2487–2498.
- Thatcher, J.D., McBride, B., Katula, K.S., 1995. Promoter binding factors regulating cyclin B transcription in the sea urchin embryo. *DNA Cell Biol.* 14, 869–881.
- Thuret, J.Y., Valay, J.G., Faye, G., Mann, C., 1996. Cdv1 (CAK in vivo), a novel Cdk-activating kinase. *Cell* 86, 565–576.
- Trooskens, G., De Beule, D., Decouttere, F., Van Crielinge, W., 2005. Phylogenetic trees: visualizing, customizing and detecting incongruence. *Bioinformatics* 21, 3801–3802.
- Voronina, E., Marzluff, W.F., Wessel, G.M., 2003. Cyclin B synthesis is required for sea urchin oocyte maturation. *Dev. Biol.* 256, 258–275.
- Wallenfang, M.R., Seydoux, G., 2002. cdk-7 is required for mRNA transcription and cell cycle progression in *Caenorhabditis elegans* embryos. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5527–5532.
- Wohlbold, L., Laroche, S., Liao, J.C., Livshits, G., Singer, J., Shokat, K.M., Fisher, R.P., 2006. The cyclin-dependent kinase (CDK) family member PNQALRE/CCRK supports cell proliferation but has no intrinsic CDK-activating kinase (CAK) activity. *Cell Cycle* 5, 546–554.
- Woo, R.A., Poon, R.Y., 2003. Cyclin-dependent kinases and S phase control in mammalian cells. *Cell Cycle* 2, 316–324.
- Ye, X., Zhu, C., Harper, J.W., 2001. A premature-termination mutation in the *Mus musculus* cyclin-dependent kinase 3 gene. *Proc. Natl. Acad. Sci. U. S. A.* 98, 1682–1686.
- Yoshida, K., 2005. Identification of a novel cell-cycle-induced MCM family protein MCM9. *Biochem. Biophys. Res. Commun.* 331, 669–674.
- Zakian, V.A., 1995. ATM-related genes: what do they tell us about functions of the human gene? *Cell* 82, 685–687.

**Development of wideband microstrip geometries, RFID tag mounting  
structures and radiated emissions testing capabilities**

Project Investigator:

Dean Arakaki  
Electrical Engineering  
California Polytechnic State University  
San Luis Obispo, California

The primary objectives of the 2005-2006 C<sup>3</sup>RP project have been to:

- Develop and apply high-impedance (EBG: electronic bandgap) ground planes through HFSS (Ansoft electromagnetic simulation software) to allow direct placement of RFID tag geometries
- Acquire fabrication capabilities to construct prototype EBG ground planes
- Develop an Electromagnetic Compatibility (EMC) chamber with radiated emissions testing capabilities
- Establish computer models of multiband fractal antennas to operate in the ISM (Industrial-Scientific-Medical) bands
- Fabricate fractal antennas and test for matching and radiation pattern performance over the ISM bands to verify conformance to simulation results
- Computer models of RFID tags mounted on flexible mounting structures

### EBG – Electronic Bandgap – Structures

High-impedance (high-Z) ground planes are useful for planar antenna geometries as they allow direct contact with the ground plane. This is due to in-phase images produced by high-Z ground planes (image theory). This is especially advantageous if the high-Z ground is low-profile; much less than  $\lambda/4$  at the operating frequencies.

To quantify the performance of the high-Z ground, a computer model was developed and reflection characteristics examined: see Fig. 1 below.

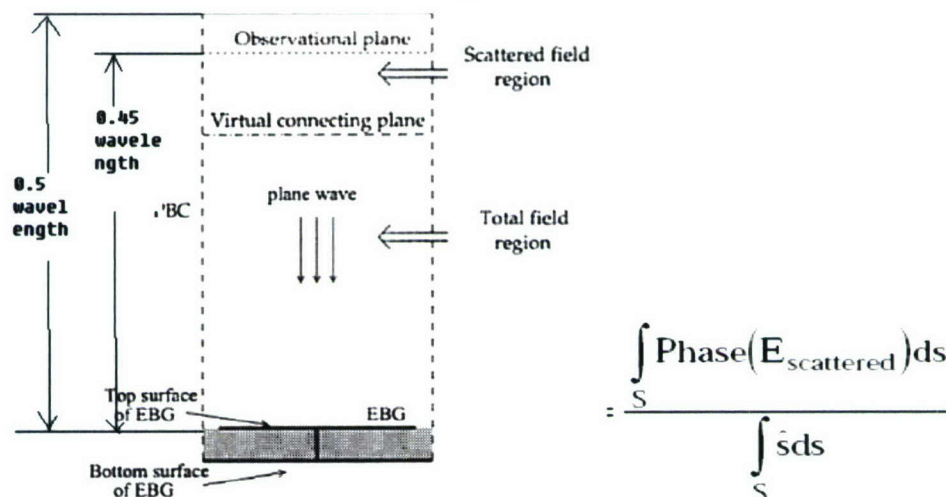
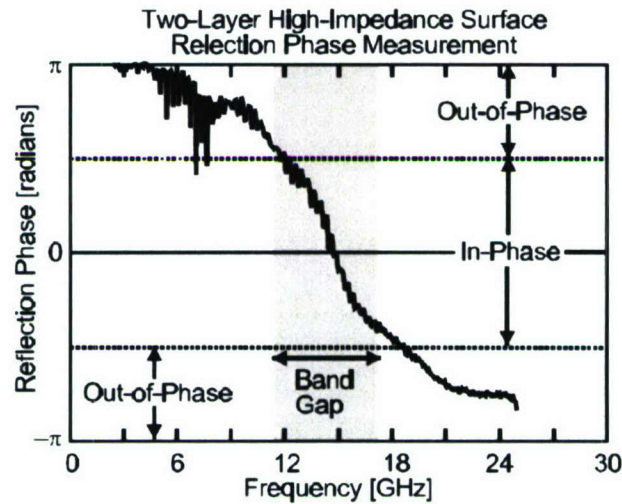


Fig. 1: Measurement Technique for Verifying EBG Performance

A uniform plane wave is established at the observational plane and used to excite the EBG structure. This wave impinges on the EBG structure, which creates a reflection. The region labeled “total field region” is so-named because it contains both incident and reflected radiation. The virtual connecting plane serves as an effective directional coupler that allows only the reflected wave to pass to the observational plane. The reflected wave

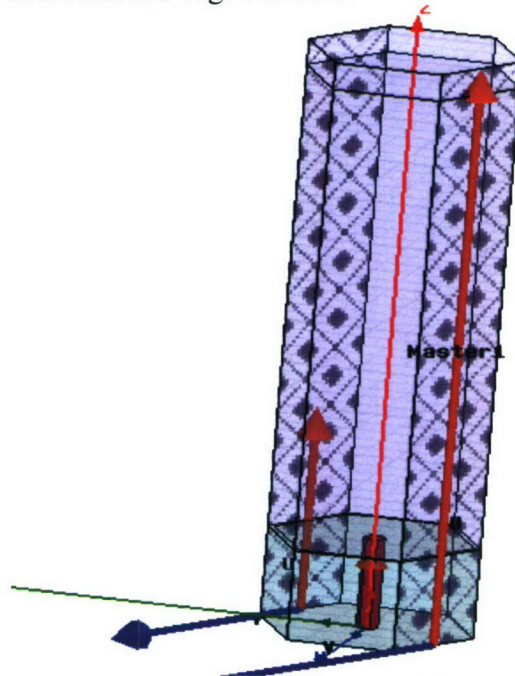
is measured at the observational plane and compared to the incident wave. The phase difference is noted at each point in the observation plane and integrated over the cross-sectional surface. This integral is divided by the total cross-sectional area to arrive at the overall averaged phase response: Fig. 1.

A high-Z surface behaves similar to free-space, except that no radiation is passed from the radiating region through the outer boundary. This results in a +1 reflection coefficient ( $0^\circ$  phase). Hence, phase responses relatively close to the  $0^\circ$  value (within  $\pm 90^\circ$ ) is the desired result [1].



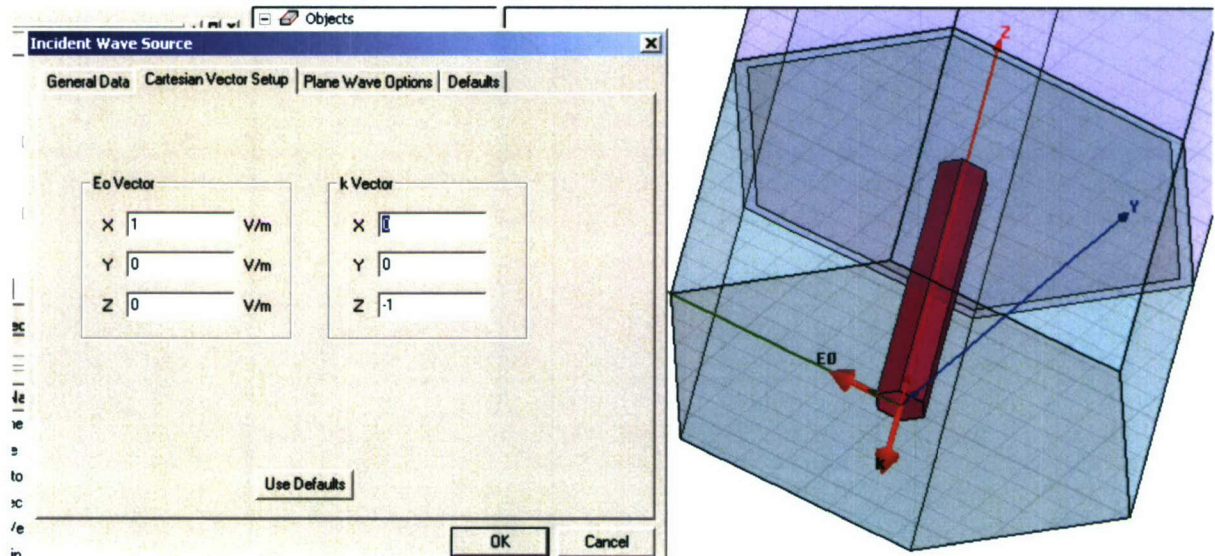
**Fig. 2:** Reflection Phase Measurement for High-Impedance Surfaces

The development of EBG structures has involved the use of hexagonally-arranged symmetry planes in HFSS: see Fig. 3 below.



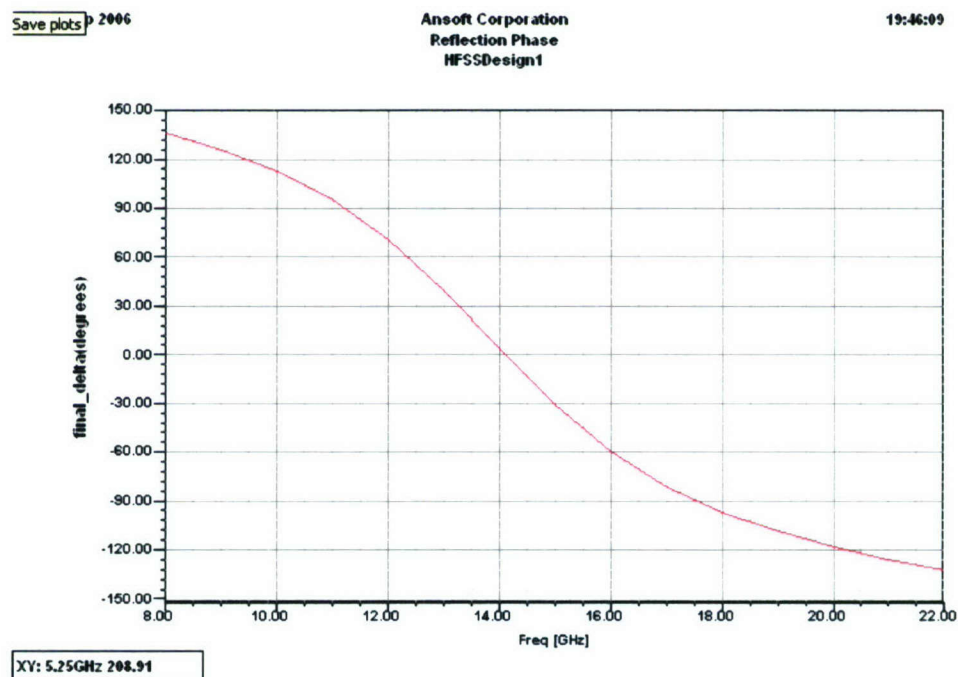
**Fig. 3:** Simulation Model for EBG Structure

The geometry at the bottom of the cell is a hexagonal version of a high-impedance structure proposed in [1]. It consists of a via hole from a perfect electrical ground up to a two-dimensional shape arranged in a honeycomb pattern. To simulate an infinite array of these structures, hexagonal symmetry planes defined around a so-called “unit cell” are employed in HFSS. The symmetry planes emulate mirrors in optical systems in the microwave frequency range (GHz).



**Fig. 4:** Expanded Views of EBG Model in HFSS

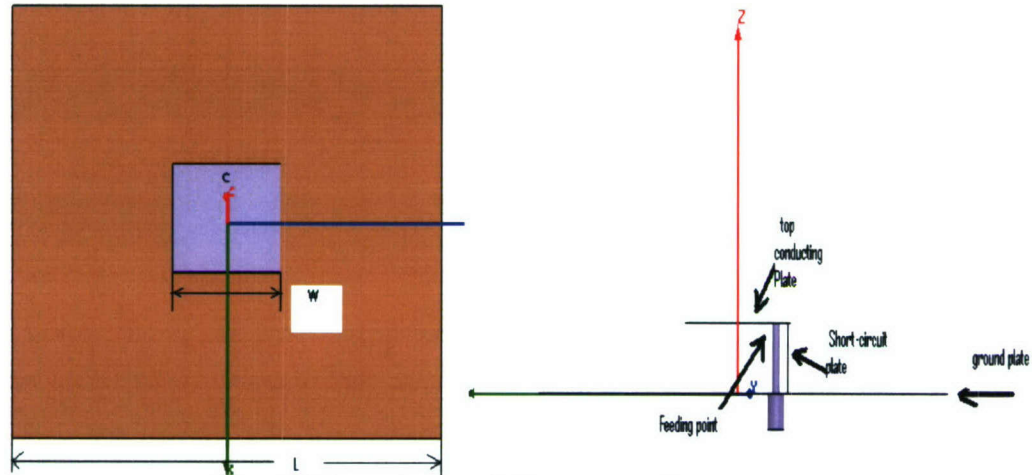
The desired response is a phase response within  $\pm 90^\circ$  in the measured reflection coefficient within the desired operating frequency. For this structure, the intended frequency region is 14GHz. HFSS simulation results are shown in Fig. 5 below.



**Fig. 5:** Reflection Coefficient Response for EBG Structure Defined in HFSS

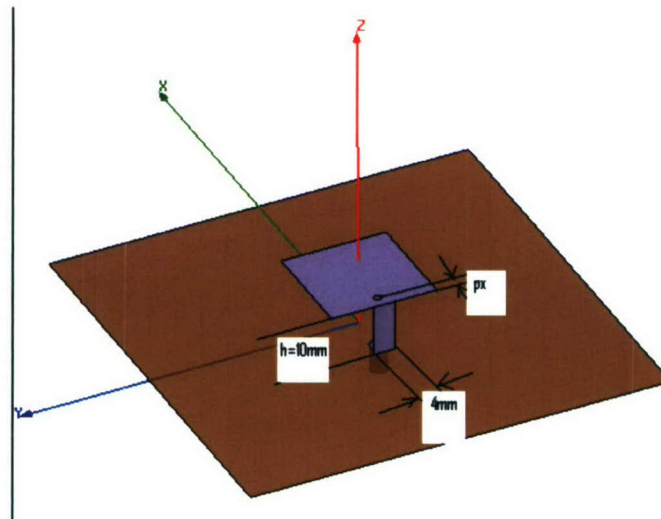
The above response indicates that the structure is operating within the  $\pm 90^\circ$  criteria from 11.25GHz to 17.5GHz.

A second structure was also examined as alternative to the EBG structure; the PIFA (planar inverted-F antenna), see Fig. 6 below.



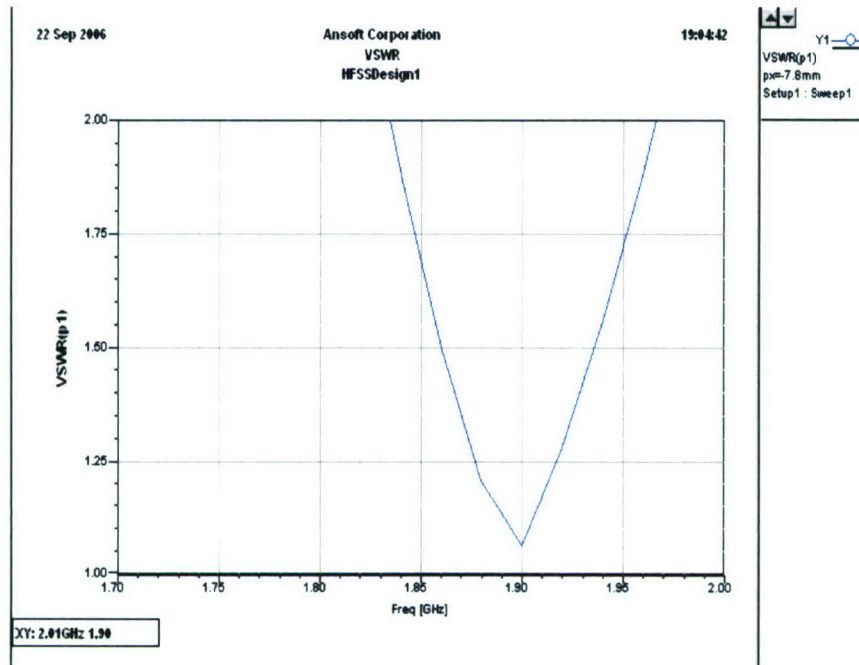
**Fig. 6:** Planar Inverted-F Antenna Geometry

This geometry has been shown to perform at a level comparable to EBG structures but with a lower profile [2]. Several dimensions on the PIFA were varied including the structure height, placement of the shorting pin, and the size of the ground plane: see Fig. 7 below.



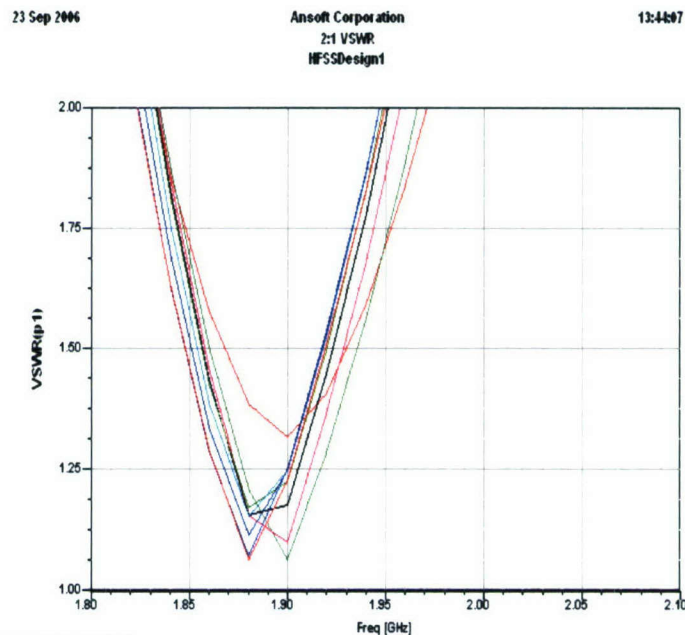
**Fig. 7:** Dimensional Variables, PIFA

A preliminary simulation of the input matching characteristics is shown in Fig. 8 below.



**Fig. 8:**  $|S_{11}|$  (dB) Frequency Response

This plot confirms a 1.9GHz resonant frequency with a VSWR = 2 bandwidth of from 1.835GHz to 1.935GHz (5.3%). Variation of the ground plane size (square of dimension L in mm) is shown below.



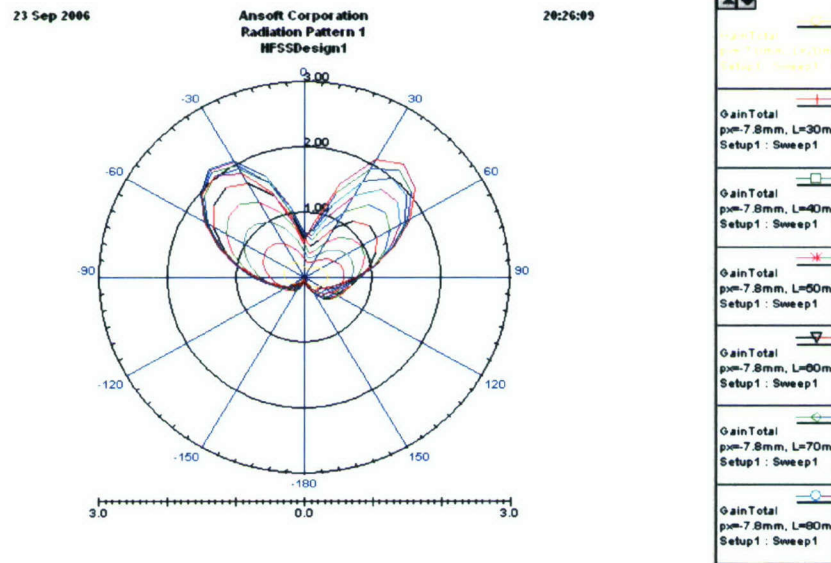
**Fig. 9:**  $|S_{11}|$  (dB) vs. Frequency for Various Ground Plane Sizes

Table 1 below summarizes the VSWR = 2 bandwidth as a function of ground plane size.

Ground plane size in wave lengths	relative bandwidth(%)
0.25	0
0.375	7.4
0.5	6.8
0.625	6.8
0.75	6.3
0.875	6.4
1	6.1
1.125	6.4
1.25	6.4
1.375	6.45

**Table 1:** Percent Bandwidth vs. Size of Ground Plane

Note that as the ground plane size increases beyond 1.125 $\lambda$ , the bandwidth remains relatively constant. Radiation patterns were also analyzed as a function of ground plane size: see Fig. 10.



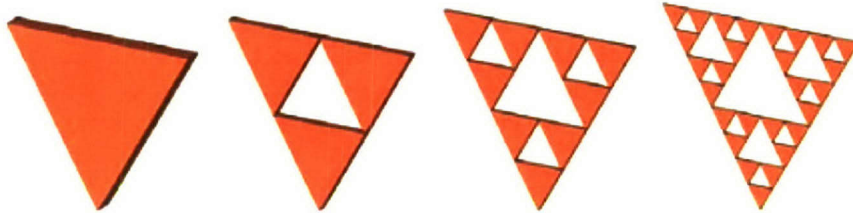
**Fig. 10:** Radiation Patterns for PIFA vs. Ground Plane Size

This plot shows that radiation patterns also reach a constant state when the ground plane reaches the 1.1 $\lambda$  size.

### Fractal Antennas

Fractal patterns are composed of self-similar structures constructed in an iterative fashion. Sierpinski Gasket pattern is one type of fractal pattern composed of triangular elements. The construction of the Sierpinski Gasket pattern involves the removal of the central triangle (generator) so that the remaining triangles, specially the bottom triangle, will have the scale desired which correspond to the frequency band spacing. The triangle

removal process is repeated for the remaining triangles and this process can be repeated to the desired number of iterations. The true mathematical definition of a fractal pattern requires that the generator is applied an infinite number of times but earlier studies [4] indicate that “higher” numbers of iterations produces intricacies that are much smaller than the wavelength of the resonant frequency. These intricacies can be eliminated to simplify the geometric pattern into what is called a “prefactal.”



**Fig. 11:** Sierpinski Gasket fractal pattern with 0.5 scale factor.

## PROJECT GOAL

The goal of this project is to construct a multi-band low profile antenna that matches the ISM band. To match at multiple frequencies, the geometric scale factor of the fractal pattern has to be adjusted. Dr. Song’s published work on the perturbed Sierpinski Gasket antenna [6] was selected since his work shows that the scale factor can be adjusted for the desired band spacing. The coplanar ground plane also gives his fractal antenna a low profile. Some modifications in Dr. Song’s gasket antenna are necessary so that it can be used in the ISM bands.

## METHOD

The ISM (industrial, scientific and medical) bands are consisting of 0.9 GHz, 1.8 GHz, 2.4 GHz and 5.8 GHz frequency bands. One possible matching scheme is to exclude the 1.8 GHz band and just match the remaining bands as shown below with their ratios:

$$1^{\text{st}} \text{ band and } 2^{\text{nd}} \text{ band ratio} = 0.9/2.4 = 0.38$$

$$2^{\text{nd}} \text{ band and } 3^{\text{rd}} \text{ band ratio} = 2.4/3.9 = 0.61$$

$$3^{\text{rd}} \text{ band and } 4^{\text{th}} \text{ band ratio} = 3.9/5.8 = 0.67$$

The fundamental frequency (0.9 GHz) and the 2<sup>nd</sup> resonant frequency (2.4 GHz) have a ratio of 0.38. Earlier published works (1) (3) on the Sierpinski gasket antenna mentioned that the first two resonant frequencies have a scale factor of 0.3 to 0.35. A third extra resonant frequency of about 3.9 GHz is introduced to achieve a scale factor of about 0.65 for matching both the 2.4 GHz and 5.8 GHz.

The initial height of the main triangle is determined by the relationship given in earlier work (1) where the perimeter of the main triangle is determined to be slightly more than half the wavelength of the fundamental frequency. Since the fundamental frequency is 0.9 GHz, the initial height of the main triangle is determined to be 48.2 mm. Furthermore,

Dr. Song mentioned in his work [6] that the matching of the fundamental frequency can be improved by the dimensions of the coplanar ground plane.

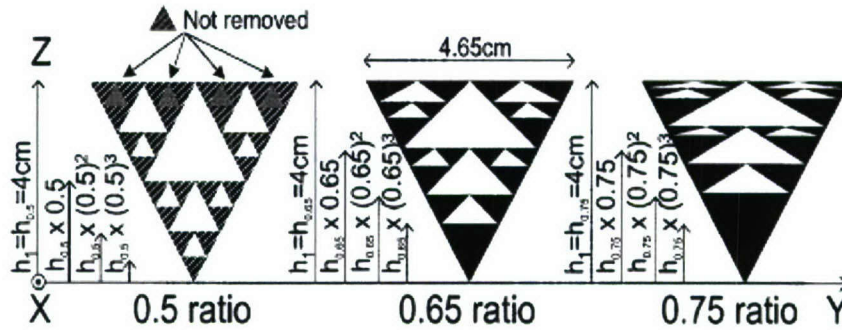


Fig. 1. Ratio of the perturbed fractal Sierpinski gasket antenna.

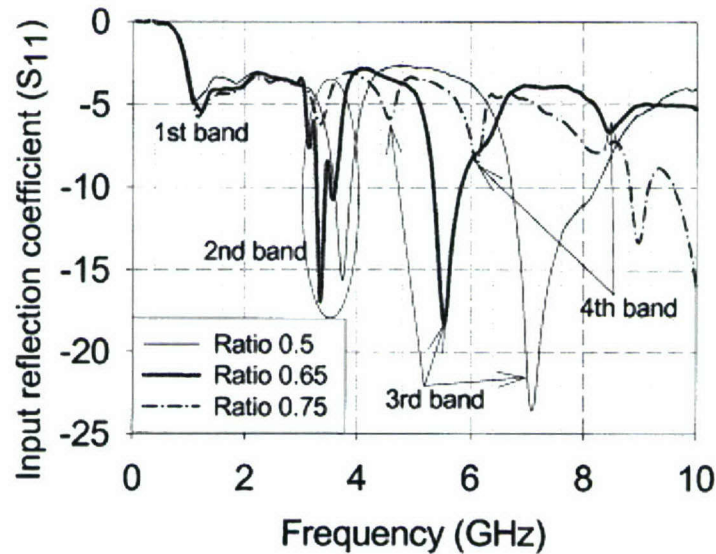


Fig. 2. Measured input reflection coefficient of the antennas shown in Fig. 1.

**Fig. 12:** Perturbed Gasket Antennas and their Reflection Coefficient [6]

The minimum trace width is limited by the fabrication process which is limited to about 0.2 mm (8 mil) in width. And then the 50 ohm feed line width is determined using ADS LineCalc. The initial ground plane dimensions are 50 mm by 100 mm.

Using basic trigonometric properties, the dimensions of the gasket antenna is parameterized as a function of the main triangle height, scale factor, flare angle, substrate thickness and trace width as shown in the Fig. 13 below. A lumped port is used for the HFSS gasket antenna model and the airbox is drawn with about  $\lambda_0/4$  height. After the parameters and constants were declared, the optimized dimensions were determined by HFSS Optimetrics for best matching of the resonant frequencies.

Using the optimized dimension, the Gerber files is created and sent out for fabrication. The fabricated gasket antenna is then tested in the Anechoic chamber for radiation pattern

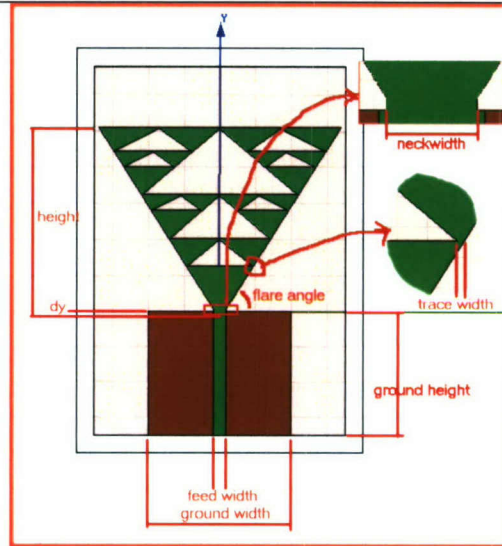
using a standard gain horn. The reflection coefficient is also measured using the HP network analyzer.

## RESULTS

**Table 2.** Optimized Dimensions.

Height = 52.8mm  
Flare = 1.003 radians (57.47 degrees)  
Scale Factor = 0.65  
Trace Width = 0.25mm  
Feedline Width = 3.5mm  
Neck Width = 3mm  
Dy = -1.95mm  
Ground Plane dimension:  
40mm X 35mm

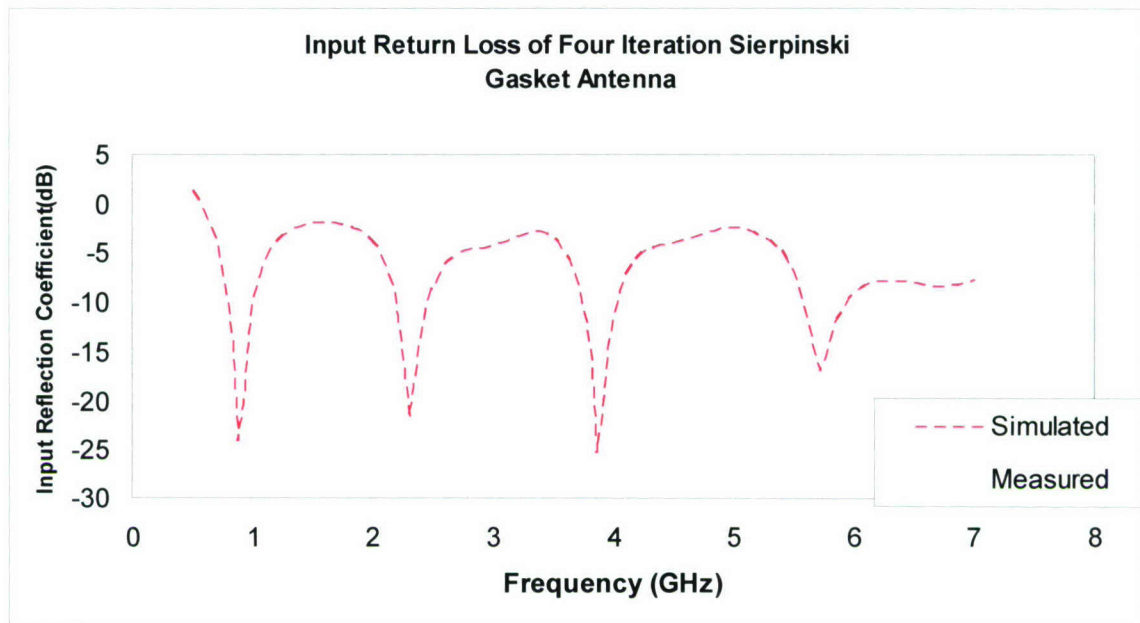
Substrate Values:  
62 mil FR4 double sided  
Permittivity, Dk: 4.4  
Dielectric Loss Tangent, Df: 0.016



**Fig. 13:** HFSS model of the perturbed Sierpinski Gasket antenna with the parametric variable dimensions



**Fig. 14:** Fabricated Sierpinski Gasket Antenna

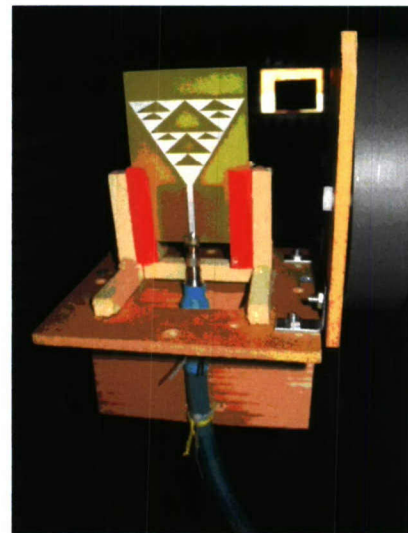


**Fig. 15:** Input Return Loss of the Sierpinski Gasket Antenna

**Table 3.** Desired vs. Measured frequencies.

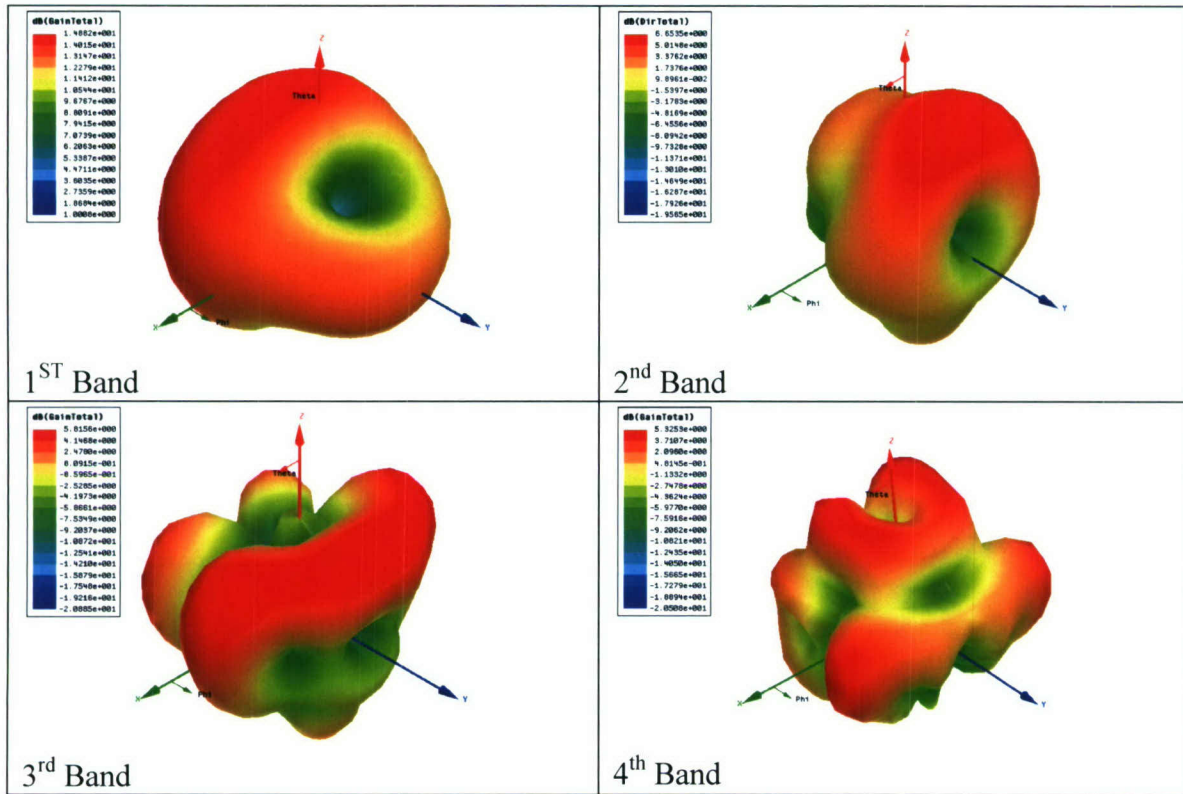
Center Frequencies			
Band	Desired	Measured	% difference
1st Band	0.9	1.19	32.22
2nd Band	2.4	2.57	7.08
3rd Band	3.9	3.97	1.79
4th Band	5.8	5.87	1.21

Band Spacing Factor (Measured)			
$f_1/f_2 =$	1.19/2.57	$=$	0.46
$f_2/f_3 =$	2.57/3.97	$=$	0.65
$f_3/f_4 =$	3.97/5.87	$=$	0.68



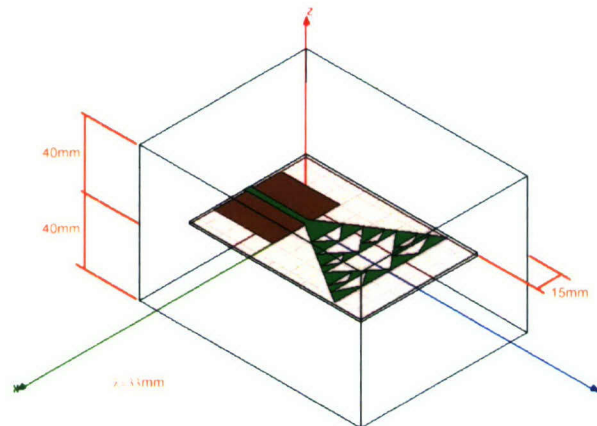
**Fig. 16:** Sierpinski Gasket Antenna in the Anechoic Chamber

The reflection coefficient measurement result shows that there is a high percent difference of more than 32% in the desired fundamental frequency to the actual measured frequency. The second resonant frequency has a closer match at 7% percent difference. The third and the fourth improves to less than 2% percent difference.

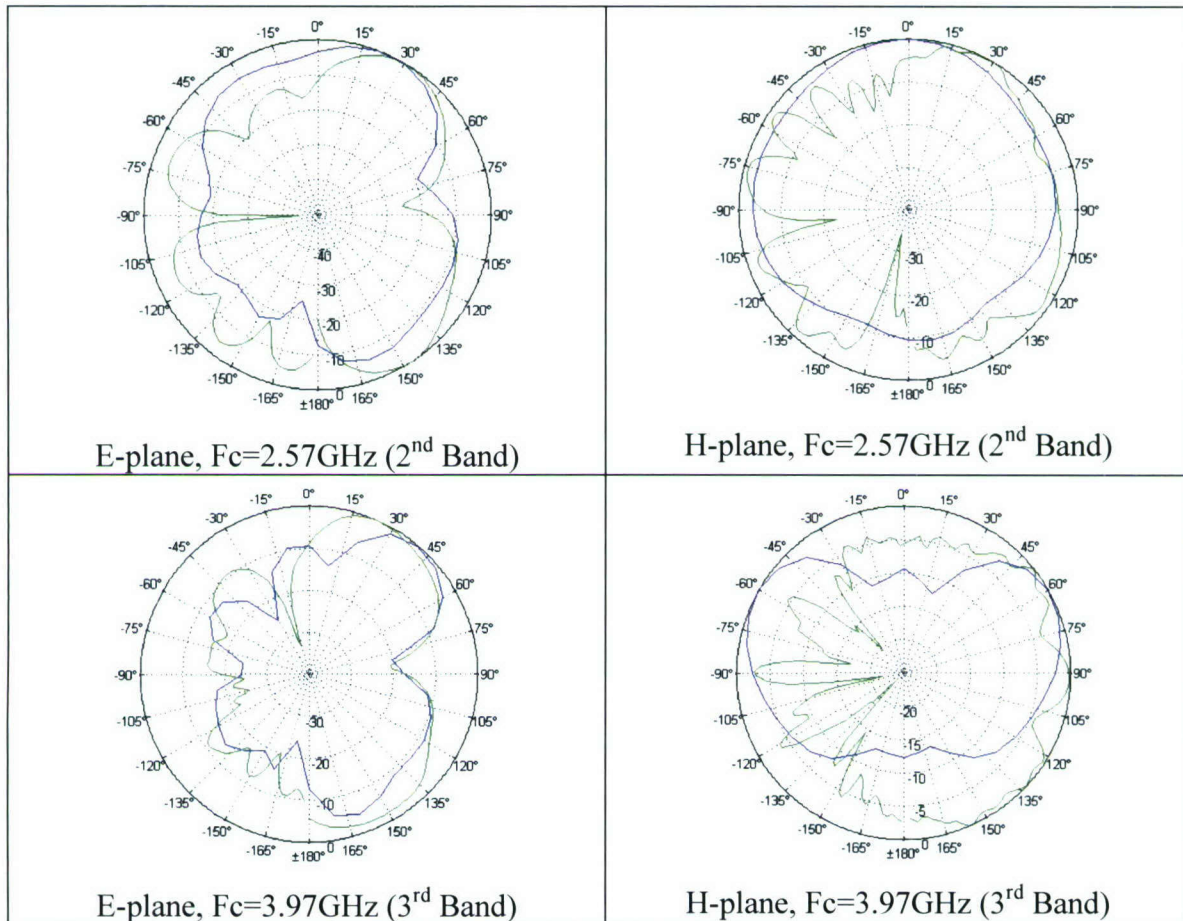


**Fig. 17:** HFSS 3D Polar Gain Plots for the Sierpinski Gasket Antenna

The 3D gain pattern in Fig.17 shows that at the fundamental frequency the radiation pattern is a simple donut shape but becomes more complex at higher frequencies. At the 1<sup>st</sup> band frequency the whole antenna is activated while at higher frequency smaller elements of the antenna are activated which result to a more complex radiation pattern.



**Fig. 18:** HFSS model of the Sierpinski Gasket with reference axis



**Fig. 19:** Normalized co-pol pattern. Blue plot - HFSS model radiation pattern.  
Green plot -Actual measurement.

The E-plane radiation pattern shown in Fig. 19 resembles the 2<sup>nd</sup> and 3<sup>rd</sup> band E-plane pattern in Dr. Song's paper. The measured H-plane pattern in figure 4 differs from the circular omni-directional pattern in the paper. The distortion on the left side of the actual H-plane measurement may be caused by the roll head blocking the antenna. It can also be observed that the H-plane radiation pattern at 2.57 GHz is almost omni-directional, like that of a dipole antenna. The orientation of the 3D plot above is based on the HFSS 3D plot so that the angle of the AMS VI plot data has to be reversed when plotted with the HFSS result. The simulated and the actual radiation pattern polar plot is superimposed using a Matlab program listed in Appendix 2.

## CONCLUSIONS – Fractal Antennas

The reflection coefficient results in Fig. 15 shows that the fundamental frequency shifted significantly but closer match is attained at higher resonant frequencies. A simulation of the gasket antenna with a larger airbox has been done but has not been fabricated as of this time.

Potential further study of the perturbed Gasket antenna can be made in phase array applications combined with EBG ground planes.

## Overall Project Conclusions

All project objectives have been met with the exceptions of the radiated emissions chamber and the study of flexible mounting structures for RFID tags. This is due to lack of student interest in these projects. The C<sup>3</sup>RP funding for these projects was redirected toward the acquisition of fabrication equipment (printed circuit board milling machine) for the production of microstrip (including fractal) antennas and RF (radio frequency) systems. Funding was also used to purchase L-band standard gain horn antennas (1.7 – 2.6GHz) to extend the anechoic chamber's testing range to include the important 2.4GHz ISM band. This allowed the radiation pattern testing of the fractal antenna at the 1.8GHz and 2.4GHz frequencies.

## References

1. D. Sievenpiper, L. Zhang, R. F. J. Broas, N.G. Alexopolous, E. Yablonovich, High-Impedance Electromagnetic Surfaces with a Forbidden Frequency Band, IEEE Trans. Microwave Theory and Techniques, Vol. 47, Issue 11, Nov. 1999, pp. 2059-2074.
2. M. N. Mollah and N. C. Karmakar, Planar PBG structures and their application to antennas, Antennas and Propagation Society International Symposium, Vol. 2, July 2001, pp. 494 – 497.
3. Analysis of Photonic Bandgap Surfaces Using Ansoft HFSS (Application Note)
4. G. F. Tsachtsiris & C. F. Soras, "Analysis of a Modified Sierpinski Gasket Monopole Antenna Printed on Dual Band Wireless Devices," IEEE Transaction on Antennas and Propagation, Vol. 52, No. 10, Oct. 2004, pp. 2571 – 2579.
5. J. P. Gianvittorio & Y. Rahmat-Amii, "Fractal Antennas: A Novel Antenna Miniaturization Technique and Applications," IEEE Antennas Propagation Magazine, Vol. 50, No. 1, Feb. 2002, pp. 20 – 36.
6. C.T. P. Song, P. S. Hall, H. Ghafouri-Shiraz, "Perturbed Sierpinski Multiband Fractal Antenna with Improved Feeding Technique," IEEE Trans. Antennas and Propagat., Vol. 51, No. 5, May 2003, 1011 – 1017.

## **Validation of RFID livestock identification systems**

Project Investigator:

Jonathon Beckett  
Animal Science  
California Polytechnic State University  
San Luis Obispo, California

*Project Summary:*

Particularly since the tragic consequences of the combination of mad cow disease outbreak and Foot and Mouth Disease epidemic in Europe during the past 15 years, the public has demanded assurance that the food supply is being regulated in a manner that assures safety in their food supply. The United States consumers have been the recipients of the safest food supply in the world. However, the perception changed on December 23, 2003, when the first case of mad cow disease was discovered in the United States. Suddenly, the foreign animal diseases (Foot and Mouth Disease, Mad Cow Disease, Avian Flu, etc.) were distinct possibilities affecting the food supply, and food animal production practices in the United States. Further, the possibility of intentionally introducing a foreign animal disease to disrupt the US food supply has certainly risen to one of the highest concerns of US Homeland Security. The United States Department of Agriculture, already planning a National Animal Identification Program, redoubled efforts and accelerated the timelines to achieve animal identification such that any animal could be traced within 48 hours. In this way, diseases could be tracked and threatened animals could be isolated before any meat enters the food supply. The National Animal Identification System is most likely to be based on radio-frequency identification ear tags that contain a unique digital number. Unfortunately, unique identification using RFID in livestock has not been validated. The current project addressed this question by tracking 1,100 cattle that transferred locations several times to a processing plant. The animal's identities are being verified through DNA fingerprinting to ensure that each animal's ID corresponds to the same animal throughout its lifetime.

*Significance:*

Recent devastating outbreaks such as foot and mouth disease, mad cow disease and avian flu plus contamination recalls and bioterrorism are driving strict new legislation on food traceability. In addition, consumers are also demanding more information about the food they consume (as do the police and customs).

Spurred on by recent food scares around the world, such as mad cow disease and bioterrorism fears, governments are forcing the adoption of food traceability systems. Although the United States has, for the most part, escaped the dramatic risks associated with mad cow disease, there have now been two confirmed cases in the US, with a third case that has been identified as inconclusive during the third week of July, 2005 (See table).

USDA Enhanced BSE Surveillance  
*# of Cattle Tested and Inconclusive Test Results*

	# of Tests	Inconclusive	Negative	Positive
Jun-04	12,398	2	2	
Jul-04	15,828			
Aug-04	24,383			
Sep-04	21,107			
Oct-04	25,476			
Nov-04	36,961	1		1
Dec-04	31,323			
Jan-05	42,002			
Feb-05	42,757			
Mar-05	52,755			
Apr-05	34,015			
May-05	30,196			
Jun-05	31,194			
Jul-05	18,422	1	?	?
Total	418,817			

In addition to the threat posed by foreign animal diseases, US Homeland Security is very concerned about the possible introduction of a disease in domestic animal species as a means of inciting fear in the public and causing massive economic chaos in the US. The value of the sales of cattle and calves in 2000 combined to an amount totaling \$40.76 billion in the United States. This accounts for more than 1/5 of all agricultural sales making the beef sector the largest single agricultural enterprise. The estimated \$40.76 billion of gross output from beef production activity supports an additional \$147.4 billion of economic output for a total of \$188.4 billion of direct and indirect economic activity throughout the U.S. economy. Just the one cow identified as having mad cow disease on

December 23, 2003, caused a loss to the cattle industry that has been estimated by one source to be up to \$4.7 billion dollars in 2004 alone (lost export market accounting for the majority of that loss). Being able to rapidly identify animals that are carrying a disease can mean the difference between a relatively inconvenient loss of livestock, versus an epidemic that has far-reaching consequences.

Traceability has become a buzzword in the food industry. Consumer demands for higher-quality foods and more variety have never been greater. Consumers have long told food producers that food safety is nonnegotiable. The recent press regarding animal diseases has shaken the confidence of consumers and thereby threatens the long-term sustainability of production.

In order to reduce the threat of spreading disease throughout the animal industries, the USDA is in the process of implementing a plan to identify all animals in the US National Animal Identification Plan (NAIP). Identification of animals will occur through either electronic (radio-frequency identification; RFID) or biometric means. The foregone conclusion is that RFID will be the predominant means of providing a unique identifier to each animal to enable food traceability.

Everyone from producer to retailer will be affected by food traceability. Although tens of millions of RFID tags have been applied to livestock and food and millions of biometric procedures have been carried out, there is no thorough assessment of the (1) readability of the tags at production speeds, (2) validation that the RFID stays with the animal throughout its life (tag retention), and (3) confirmation that the RFID unique number is consistent throughout the animal's life.

All domestic animal industries are pursuing the traceback mandates independently. There are strong incentives, aside from the governmental regulations. Several major food suppliers and retailers, including McDonalds and Wal-Mart are mandating full traceability from suppliers for raw materials and products. This is a prelude to tagging everything. It has legal push with the new European Union legislation in 2005 demanding "one up one down" traceability and the US Homeland Security legislation demanding unprecedented levels of traceability.

A plethora of work has been conducted to ensure the accuracy of RFID technology, particularly with retailers such as Wal-mart expected to completely switch inventory and marketing identification from barcode scanning to RFID. All facets of agricultural and nonagricultural industries are implementing RFID into packaging and transportation systems. The technology works and has tremendous potential to revolutionize inventory control and sales. Indeed, RFID has been used for approximately 5 years in the animal industries. The pet food industries have utilized RFID as animal identification by placing a subcutaneous chip in dogs and cats (Home Again). These chips can be read by veterinarians, animal control and animal shelters to ensure that animals are returned to their owners. The domestic animal species are utilizing ear tags implemented with an RFID chip. The domestic animal species are employing an external RFID chip (ear tag)

rather than an internal chip (subcutaneous) to eliminate the risk of chip migration in the body into areas that might enter the human food supply.

Unfortunately, the uniqueness of the RFID technology has not been fully assessed. In addition, the retention rates, “read” rates and traceability have not been fully measured. In fact, an email was just received this past week documenting the poor read rates (up to 40% failure rates) in an Australian sales barn. This communication provides little confidence that the system is ready to go online. In order to determine the uniqueness of RFID technology, the RFID identifier must be measured against unique biometric data from each individual animal. The biometric data selected in for the proposed research is a set of 12 micro-satellite markers that can be used to fingerprint the animal for identification. This strategy has the potential to confirm, without question, that RFID will provide lifetime unique identification that can be used in a nationwide traceback system.

How does it relate to the technology park initiative and areas of focus or emphasis? The current research ultimately addresses one of the major concerns of US Homeland Security: Preserving a safe food supply and minimizing risk of biosecurity challenges by terrorists. The research will clearly dovetail with the campus-wide effort to develop a center of excellence in RFID technology. Further, when the RFID technology in cattle is validated, Cal Poly would be poised to establish a leadership role in developing a database that would serve the industry.

*Detailed description of work:*

Project objectives:

- 1) Establish rate of identity uniqueness
- 2) Determine the readability of tags
- 3) Determine retention rate of RFID tags during lifetime of animals

Cattle were identified with RFID eartags at two different ranches of origin. The first, Kahua Ranch, Kamuela, HI, has a long-standing relationship with Cal Poly in which they send cattle (550 head) to the Swanton Pacific Ranch each January. Cattle graze during the grass season and are shipped to a feedlot in Idaho in July. The second ranch is a local ranch in Santa Margarita (Santa Margarita Ranch). In their typical production system, the cattle are sent to Harris Ranch for feeding, and processed at the Harris Ranch packing plant in Selma, CA.

The two sources of cattle represent two completely different production systems. The Kahua Ranch cattle will be transported from Kamuela, HI, by boat to Oakland, CA, to a back grounding yard in Los Banos, CA, for 1 week and finally to Swanton Pacific Ranch for 6 months. Following the grazing season, the cattle were shipped to a feed yard in Bruneau, ID. Therefore, these cattle represent a very extreme handling and shipping situation. The Santa Margarita Ranch cattle, on the other hand, were RFID tagged at branding time (4-5 months of age), and shipped directly to Harris Feedlot in Coalinga, CA. The cattle remained on feed until harvest at Harris Ranch Processing in Selma, CA.

All cattle were identified at the ranch of origin. RFID tags were read using an AgInfolink Bluetooth Tag Tracker reader into a laptop computer. All records were entered into the Northwest Pilot Project (NPP) and Holstein Association's Farm Animal Identification Records (FAIR) program, both of which are USDA-certified databases to test the feasibility of tracking animals using RFID tags. DNA samples were collected via Typifix DNA collecting ear tags (supplied by Igenity). The DNA-collecting tags were stored until the animals were sold to a packing plant. At each transportation step (i.e. when cattle were shipped to a new location), the RFID tags were read and data sent to the databases.

When the animals finally reached the processing plant, DNA samples were collected and RFID tags were read. Twenty percent of all DNA samples will be sent to Merial (Duluth, GA) to be matched using their IGENITY DNATRACE™ test. The DNA data will be compared with the RFID data to ensure that RFID data are 100% in agreement with DNA fingerprinting.

To determine readability of tags (objective 2), all cattle tagged were totaled. For each read event (transportation or processing time), the number of tags read were divided by the number of cattle.

Readability = Tags read/number of cattle processed or transported.

To determine tag retention (objective 3), all cattle tagged were totaled. At the conclusion of the study, all RFID tags were collected and counted.

Retention rate = 
$$\frac{\text{tags collected}}{\text{(Number of cattle Initially tagged-dead or missing cattle)}}$$

#### *Results:*

Calves from two ranches were tagged with either half-duplex or full-duplex tags. The tags were applied while the calves were at the ranch of origin. As the calves were tagged, sex and visual ID, if available, were recorded. Coincident with the RFID tags, DNA samples were collected from the calves using new DNA-collecting ear tags. The DNA samples were collected, cross referenced with RFID tags and stored for further evaluation. The two ranches of origin were (1) Kahua Ranch, Kamuela, HI (representing a "complex" system), and (2) Santa Margarita Ranch, Santa Margarita, CA (representing a more straightforward, conventional system). As the animals moved from the ranch of origin to secondary and tertiary locations, the animal movements were collected by scanning RFID tags, and data were recorded in databases. Briefly, 495 calves from Kahua Ranch were tagged. The cattle were transported via boat to California, and grazed for 7 months at Cal Poly's Swanton Pacific Ranch. The tags were read each time the cattle were processed and again prior to the cattle being shipped to Bruneau, ID, to a feed yard. Upon arrival, the cattle were again scanned.

At the Santa Margarita Ranch, 616 calves were tagged, and DNA samples collected. The majority of the cattle were shipped to Harris Ranch Feed yard in Coalinga, CA. The cattle were scanned upon arrival to the feed yard.

In total, 1,111 calves have been tagged. The tags were either Allflex half-duplex or full-duplex, or Destron full-duplex tags. After 8 months, only 4 tags were lost out of the 1,111. All of the tags have retained their readability. The cattle will continue to be tracked, and at the harvest facility, tags will be scanned and DNA samples will be collected. Approximately 20% of the samples will be randomly selected for DNA fingerprinting to verify unique identification by the RFID tags.

As an additional validation step, all of the records have been submitted to 4 separate databases. One was developed and is maintained by Cal Poly. The other three are commercially available. The ease in using the databases and ability to trace animals is being assessed. Of the three commercially available databases, one has not worked, and efforts to continue with that program have ceased. The other two commercially available sites are maintaining the records.

The project took advantage of working relationships that Cal Poly has forged with outside businesses. At least 2 ranches supplying cattle, 2 feed yards, 2 packing plants, 2 RFID tag suppliers and 1 DNA testing company were instrumental in the success of the project.

#### *Traceability:*

The National Animal Identification System provides various chances for technological error and potential for misleading cattle producers on technology, in particular RFID ear tags. Initially, several companies developed RFID ear tags to get their piece of the pie, and as with all technology the first round of products had room for improvement. Many of the “underdog” tag companies have fallen out of the race for financial reasons and the failure of first round RFID tag design. The two leaders for RFID tags are Allflex USA, and Destron or Digital Angel Corporation. Both companies are well established in the ear tag and animal identification market, and have created the best RFID ear tags to date.

To aid producers in selecting a reliable tag, the Cal Poly animal identification study focused on the retention and readability of Radio Frequency Identification (RFID) ear tags. Additional observations on ease of application, and overall time for using electronic ID (eID) were noted to illustrate the practicality of asking producers to add electronic Identification to their identification routines.

To account for lost or damaged tags DNA samples were taken while the RFID tags were applied at the ranch of origin. A second DNA sample was taken at harvest to ensure that the animal was indeed the same animal that was tagged at the ranch of origin. In order to efficiently take the DNA sample and apply the RFID tags at the ranch of origin, the project linked an RFID tag with a DNA tag to an index number in an MS Excel spreadsheet.

A total of 1,123 head of cattle have been tagged since January 1, 2006 from two ranches, the Santa Margarita Ranch located on the central coast of California and the Kahua Ranch located in Hawaii.

On January 4, 2006, 510 pairs of tags (RFID & DNA) were applied to cattle at the Kahua Ranch on the big island of Hawaii. The Kahua Ranch is located on the western slope of the Kohala Mountains, 3000 feet above sea level. It encompasses 8,500 acres of private ranch land, and participates in many enterprises from grazing livestock; cattle, sheep and horses to Agri-tourism activities such as ATV rides and Horseback riding. The family has other ranches that total 22,000 acres of grazing land where they run 3,300 mother cows. Fortunately, the ranch already had an identification system for herd management, so the index number was linked to the visual 4 or 5 digit ranch identification number for further identification accuracy.

Hawaiian cattle producers face a unique hurdle when it comes to selling their calves; the miles of ocean separating them from the mainland. The cattle can travel either by air or sea, and this year the cattle came over to Santa Cruz, CA on a boat. The transport across the ocean took approximately 4½ days and the opportunity for tag loss rises in confinement. The animals were scanned approximately 14 days after arrival on the mainland. At that point, retention was 100% and readability was 100%. On August 4, 466 head were shipped to Bruneau ID, 31 were shipped to auction at 101 Livestock Market in Armoas, CA, and one steer was found in October still grazing at Swanton Pacific Ranch in Santa Cruz, CA.

On August 17<sup>th</sup>, 2006 the Kahua cattle were processed in Bruneau, ID. The cattle had been in Idaho 12 days since arriving from Swanton (Santa Cruz, CA). All 466 animals were scanned into the facility and only one animal came through the chute without its RFID tag. The cattle are being fed at Davis Cattle Company. Davis Cattle Company is a family owned and operated feedlot (7,500 head capacity) and cow-calf operations. The cattle were fed approximately 5 months, and were harvested in Pasco, WA, at a Tyson Foods Packing plant. Tags were scanned and DNA samples were collected.

A total of 591 calves were tagged at the Santa Margarita Ranch in Santa Margarita, CA. Differing from the Kahua Ranch, the Santa Margarita Ranch did not have a herd management ID system so the index number was linked to the gender of the calf (either a steer (s) or a heifer (h)). The Santa Margarita Ranch runs 700 cows on 14,000 acres. A typical California commercial cattle ranch, they calve during the fall and sell calves in the spring. The ranch participates in the Harris Ranch, and the calves traveled 100 miles from the home ranch to begin feeding. A total of 363 calves were shipped out of the Santa Margarita ranch in July, 2006. The cattle were shipped on trucks to Harris Feeding Company in Coalinga, CA where feed yard employees scanned the RFID tags upon arrival. The process of being loaded and unloaded on a truck increases the chance for tags to get hung up on something and come out of the ear. After the truck ride there were two animals that did not have an RFID tag and one steer whose RFID tag did not read. The three were re-tagged with a new RFID number. The tag that would not read was an Allflex Half-Duplex (HDX) tag, and had not scanned at weaning time (therefore failure

occurred before weaning). One of the lost tags was an Allflex Full-Duplex (FDX) tag which was determined by the DNA tag number that was still in the animal's ear. The other lost tag was missing at weaning time, again the DNA tag reference was used to identify the animal. The same animal had lost its DNA tag when it arrived at Harris Ranch and has no original identification devices. Overall, the retention rate from initial tagging to entering the Harris feedlot was 99.45% and readability was 99.72 %. Also noted was 6 animals had lost their DNA tags prior to shipment to Harris Ranch.

The cattle originating from Santa Margarita ranch are currently being fed at Harris Feeding Company, and are scheduled for harvest February, 2007. As the cattle are shipped to the packing plant, the goal of the project is to scan all the cattle out of the feedlot and again scan them into the packing facilities. In addition to scanning the animals, 20% of the carcasses will have DNA samples taken to verify that there was no tampering of tags.

#### *Data Tracking:*

Northwest Pilot Project, NWPP, is a USDA funded prototype for the United States Animal Identification Plan (USAIP). The NWPP was one of the first projects to develop a process that allows a livestock animal to be traced through the various industry segments and back to its origin quickly and efficiently. The NWPP system is relatively easy to use and simple to submit back to their website. Submissions are based on events such as ID tags being applied and movements to and from different premises. Once logged in on the Northwest site, there is an Excel link that allows one to open a spreadsheet with heading columns for all the necessary information to comply with the plan. These headings include the animals' ID number, premise location(s), the event taking place, etc. A producer submits data either directly or via a spreadsheet conduit. Once submitted, it is possible to view your transactions by individual animal ID or as a large group.

The second web application tested during the project is the National F.A.I.R. Program. Overall, the purpose of the two programs is similar, but the means of submitting the information is very different. While NWPP has a simple excel sheet to follow and links to upload the information on your own, F.A.I.R. does not. In order to submit information for tracking, one must be in contact with a representative of the program. Information is electronically sent via e-mail to a FAIR representative, who then uploaded the event and animals involved. Once information was submitted, it was possible to track animals by ID number or premise location. If searching by location, the database would return a list of all animals recorded to that location at the present time. The process of submitting my information was still simple, but not as quick as with NWPP. It took time (a week or two) for the FAIR representative to get information online and available for viewing.

A third vendor's software was unsuccessful; RMS LivestockTrust Software. As with the other two programs, Research Management Systems USA developed LivestockTrust to help certain states fulfill the requirements of the USAIP and satisfy the need for source verification and traceability. However, this system is not setup for producers operating

out of California. For this reason, it was not possible to register our premises or verify animal movements. No information was ever tracked via RMS LivestockTrust.

Overall, NWPP was the easiest to use. With minimal knowledge of Microsoft Excel, it was possible to organize data, download the format, and upload my animal information within 10 minutes. Animal movements and events were then immediately viewable. Submitting animal information to FAIR was still fairly simple because it only required sending information to a FAIR representative. The drawback to this method, however, was the time lapse between submission and the possibility of viewing it online. Lastly, RMS LivestockTrust did not work for a California operation, so testing of this product was rapidly abandoned.

*To be completed:*

All cattle will be harvested by the middle of February, 2007. DNA samples have been collected on all of the carcasses to date, and RFID tags have been read. To date, both retention rate and readability exceed 99% (1-year measurement). DNA verification of identity will be completed by late spring 2007.

*Description of how findings will be used:*

The scientific results will be presented at national and regional meetings including the American Society of Animal Science (ASAS) annual meeting and the Western Section meeting of ASAS. The results of the proposed study will be promoted through literature available to California and US cattle producers. The primary route of dissemination will be via publication in trade journals such as Drovers Journal and Beef magazine. The progress on the project will be promoted at Cal Poly's annual Beef Field Day. Finally, the results will be available on the principal investigator's website as a fact sheet presented in a .PDF format.

The funding of this study should be considered seed money to promote a line of research that will continue to enhance food safety for consumers and aid in the production of livestock through recordkeeping. The proposed research clearly builds upon the success of previous research at Cal Poly, and most importantly, applies current technology to increase productivity of our agricultural resources. Required investment in the technology will be calculated as an aid for producers, and return on investment will be determined for each producer interested.

*Other support:*

The C3RP funding was used to fund a study that has been used as a springboard to apply for federal funding. The USDA recently funded the CDFA with a grant to develop an economic cost analysis of RFID implementation in the California Livestock Industries. Data from the C3RP were used as preliminary data to support the grant. Cal Poly is serving as one of 3 universities (UC Davis and CSU Chico) participating on the study. The total grant by USDA was \$350,000, and Cal Poly's work comprises approximately \$83,000 of that total.

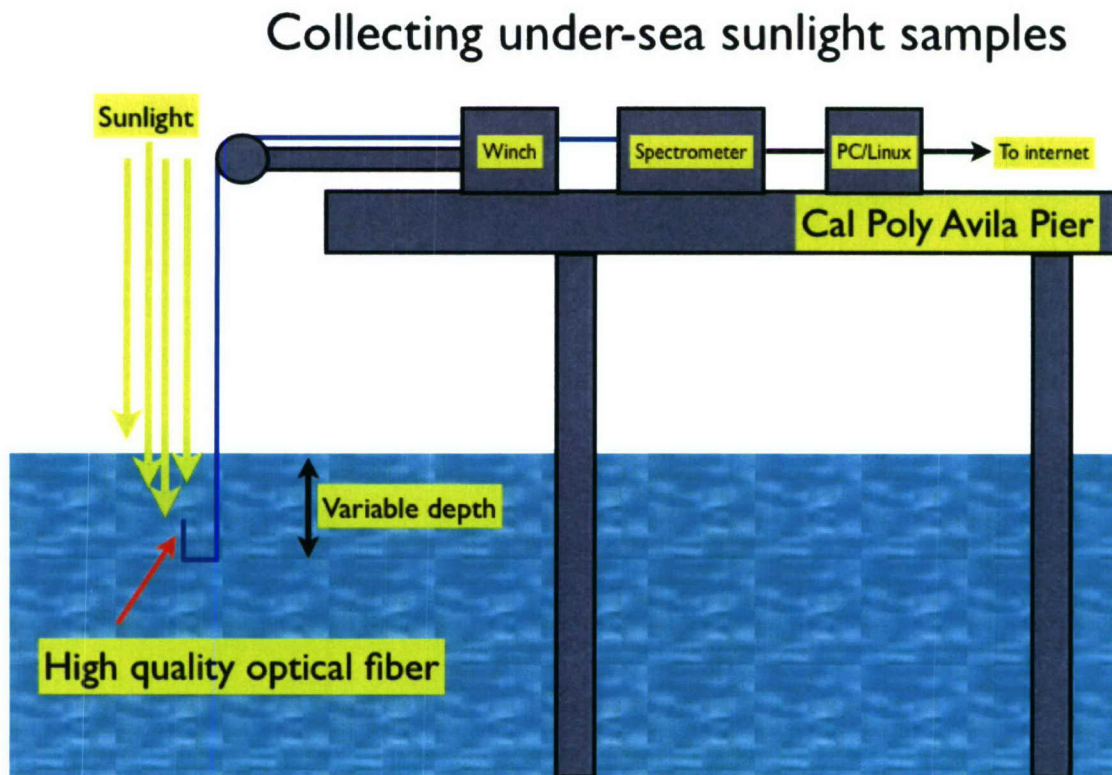
The principal investigator has also submitted a notice of intent to apply for Agriculture Research Initiative funding for 2007-08. The proposals are due February 23, 2007. Funding will support continued research in both development and application of RFID technology to identify livestock.

## **Continuous optical monitoring of a near-shore sea-water column**

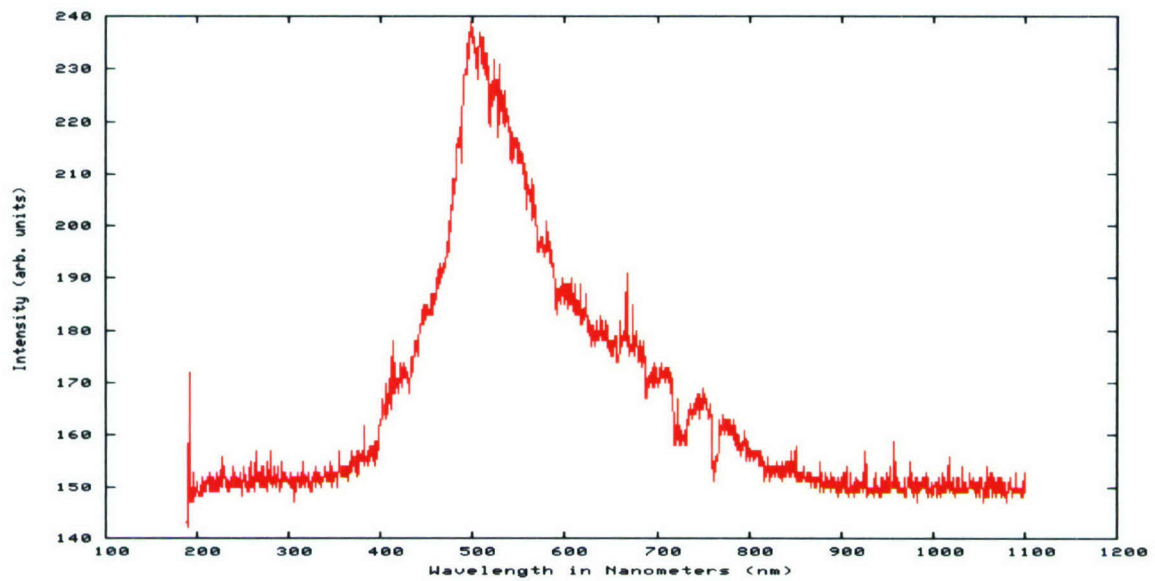
Project Investigator:

Tom Bensky  
Physics  
California Polytechnic State University  
San Luis Obispo, California

This report discusses the results from a project centered around the construction of an optical system that collects light samples in ocean water, from the surface to a depth of 30 feet. The system is housed at the Cal Poly pier in Avila Bay. In a nutshell, a computer-controlled winch system slowly lowers a high-quality fiber optic cable into the water at regular intervals (2 times per hour). The input aperture of the fiber points straight up. As the fiber is lowered, a spectrum analyzer on the pier surface receives light that entered the fiber aperture, and reports on the light intensity found at wavelengths between 200 and 1100 nm. A schematic for this system is shown here.

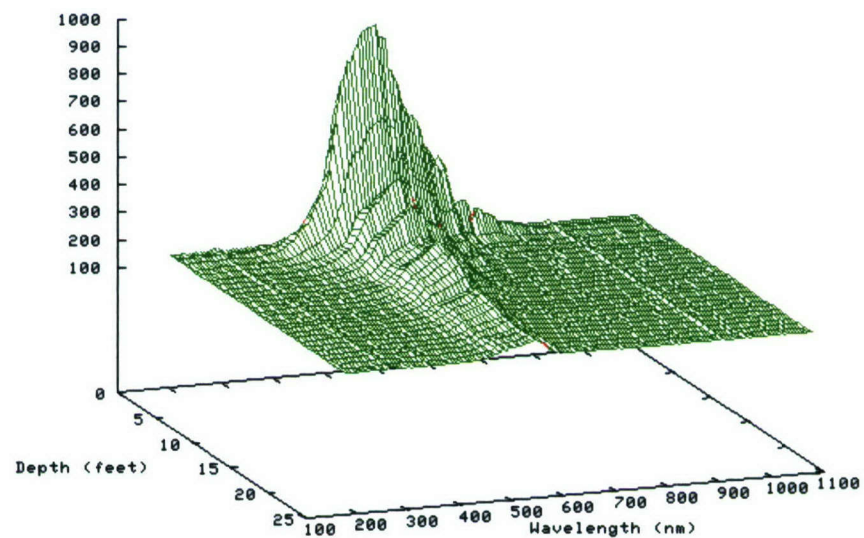


A sample spectrum taken at a given depth, is shown here.



One spectrum resembling the one above, is taken at each of 60 increasing depths, during one 0 to 30 foot "dive" of the instrument. This acquisition take approximately 70 seconds to complete. All told, the raw data from a single dive is shown here.

Intensity (arb. units)



Such a data represents a "raw data set" for this instrument, and shows the the primary purpose of this project: collecting sunlight samples under the ocean surface.

To minimize the effects of ocean fouling, the instrument remains in the water only during data acquisition runs (approximately 2 minutes at a time). All other times, it is retracted to the level of the pier, approximately 25 feet above the water. Minimizing the time the instrument spends in the sea water dramatically decreases fouling and the probability of theft or damage to the instrument. We expect this design to allow us to continue under-sea light collection studies well beyond the term of this grant. To date, we have completed the goals set out in the proposal for this work and the optical monitoring apparatus is fully operational.

The sample data set shown above is dense with information. The wavelength of interest is across the horizontal axis, between 200 - 1100 nm. The axis that extends into the page is the depth at which the spectrum was taken, and the vertical axis the relative light intensity. Such a data set is taken twice an hour between 6am and 8pm everyday. We are currently working on extracting more features from the data. As a first example, integrating under any curve, over the full wavelength rage, at constant depth, give us a single number proportional to the total number of photons that existed at that depth. Second, taking a "slice" in the data from 0 to 30 feet, at constant wavelength, gives us the attenuation of a particular wavelength of light as a function of depth. This leads directly to attenuation coefficients for a given wavelength, a key apparent property of ocean water. Samples of such analysis is shown below.

### **Deployment**

Here are two figures illustrating the deployment (left) and operation (right) of this system.

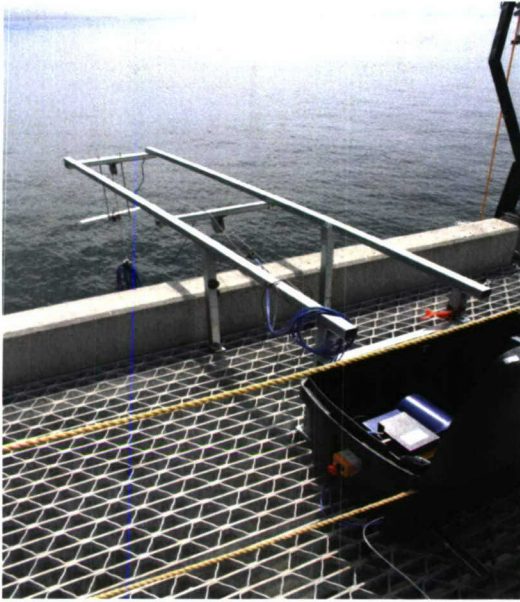


Figure 1: Boom assembly. The blue cables are the fiber optic cables that collect light from under the ocean. The winch and control electronics are in the black, weather-proof box (the lid is open for the figure).

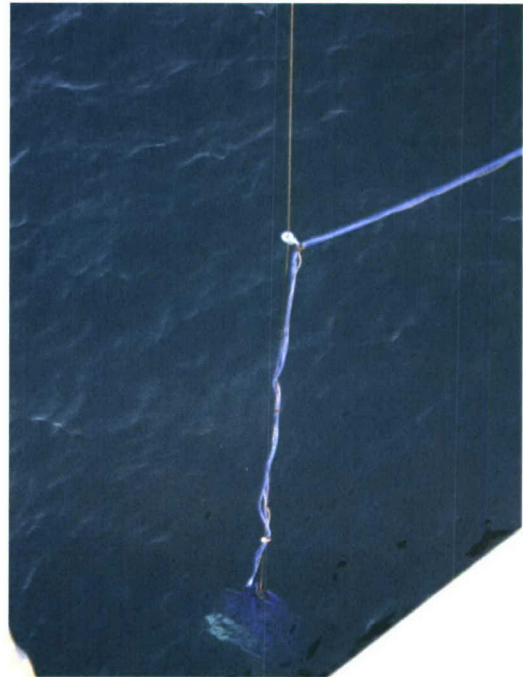


Figure 2: The fiber aperture slowly descending under the ocean surface.

### Project Status

As mentioned, this project is fully operational and has resulted in the completion of senior projects by physics majors Brenton Campbell and Bryan Neff. Tasks completed are shown in this table:

- Deployment of the boom at the Cal Poly Avila Pier
- Acquisition of a winch that will lower the instrument, through the boom's cable/pulley system, slowly into the water.
- Interfacing of the winch to a computer for fully automate winch control. This included adding a magnetic/hall switch encoder to the winch for position control to 1 inch. Computer control of the winch will allow for unsupervised data acquisition around-the-clock. We plan on programming the computer to run the winch, which will slowly lower the instrument to the 30 foot ocean bottom 4 or 5 times per hour.
- Custom Linux software to acquire a light spectra from an Ocean Optics, HR2000 spectrometer, acquired under the previous grant.
- Electronic interlocks on the boom to prevent a "runaway" event and automatic shutdown of the winch in the event it gets out of control.
- An electronic "home" sensor that tells the computer when the instrument has been pulled to its highest possible position. A strong magnet mounted on the tether triggers a reed-switch attached to the boom. The computer can sense closure of the reed switch and

shuts off the winch, assuming the instrument is "home."

- An electronic "splash" sensor which will tell the computer when the instrument first hits the ocean water on its way down, indicating a depth of 0 ft. This will give good knowledge of instrument depth, in light of the 6 foot tides in Avila Bay.
- Deploying the computer controlled winch alongside the boom at the pier. Integrating the electronics, power supply, and computer into the pier infrastructure.
- Integration of software and hardware to safely leave the apparatus unattended to acquire data several times per hour (custom 2,000 line C-program).

### **Apparatus Use**

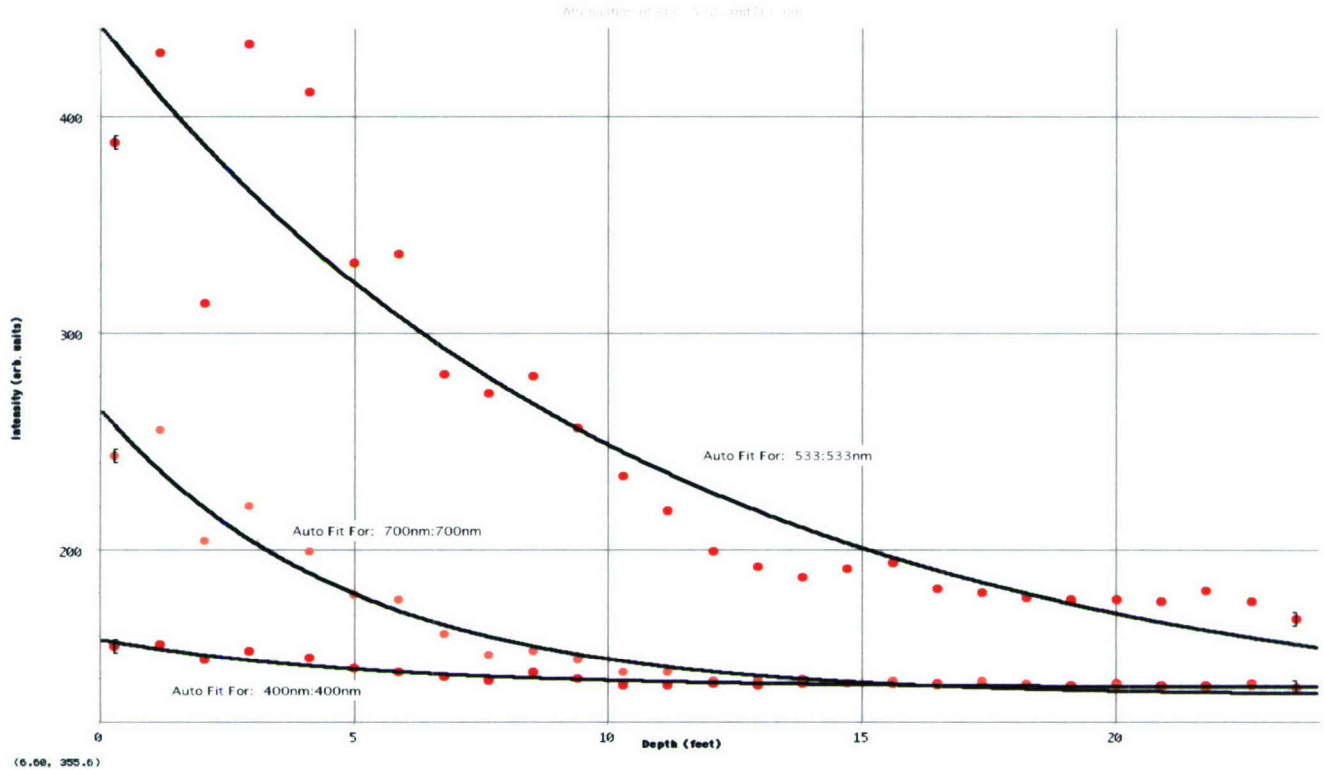
Over the next *several years*, we plan an ongoing data logging and analysis procedure, including:

- Monitoring the time-dynamics of individual wavelengths, over short periods (daily) to longer periods (months). Of particular interest will be the later summer months, when much of the water surrounding the pier turns reddish-brown. This is due to blooms of algae which is common for this time of the year. We will look carefully for optical signatures of the oncoming bloom. The camera will also acquiring images of the water when data is taken, and this should help to correlate optical signals with actual ocean surface visuals.
- Monitoring bottom reflecting using the downward pointing fiber. The percentage of downwelling light reflecting from the ocean floor is a critical parameter in optical bathymetry measurements done in many LIDAR studies. We hope to quantify this parameter measurement with various ocean conditions.
- Ocean surface reflection will be performed with another downward pointing fiber. This will yield a spectra of light reflecting from the ocean surface and can be readily compared with the light spectra incident from the sun.

### **Data Analysis**

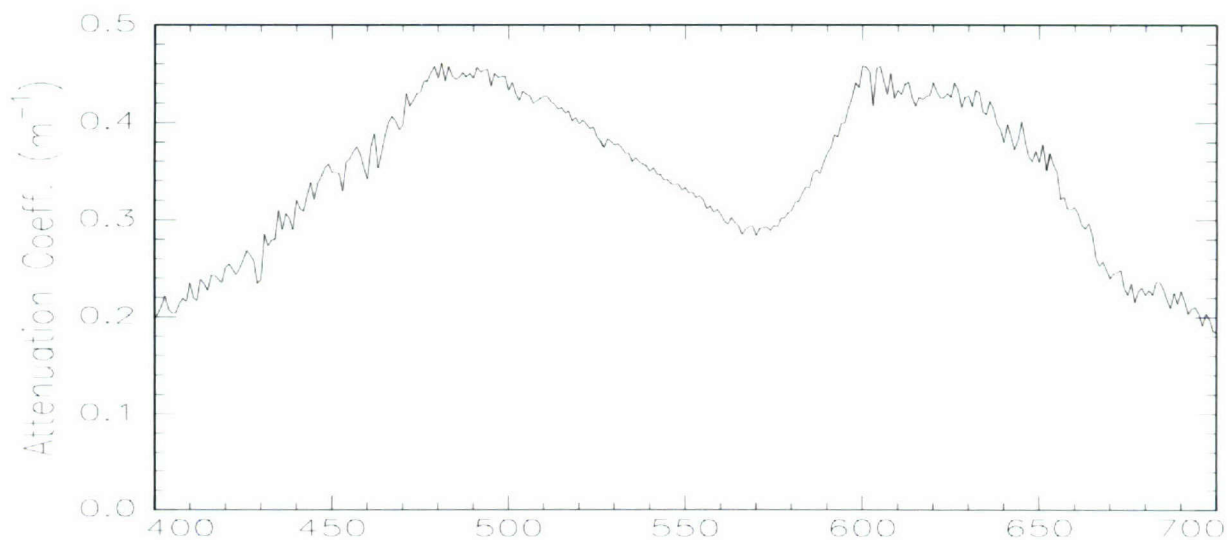
Work is on-going with this project, primarily in the area of data analysis. We are currently focusing attention on two analysis/data reduction efforts, all derived from the sample "raw data" shown above.

**I. Attenuation coefficients.** We can take "slices" of the raw data, involving intensity versus depth at constant wavelength. Data such as this reveals how a particular wavelength gets attenuated with depth as it travels into the ocean water. The so called "attenuation coefficients" remain a standard concern for any ocean-optics related study. Three such wavelengths are extracted and shown on this plot.



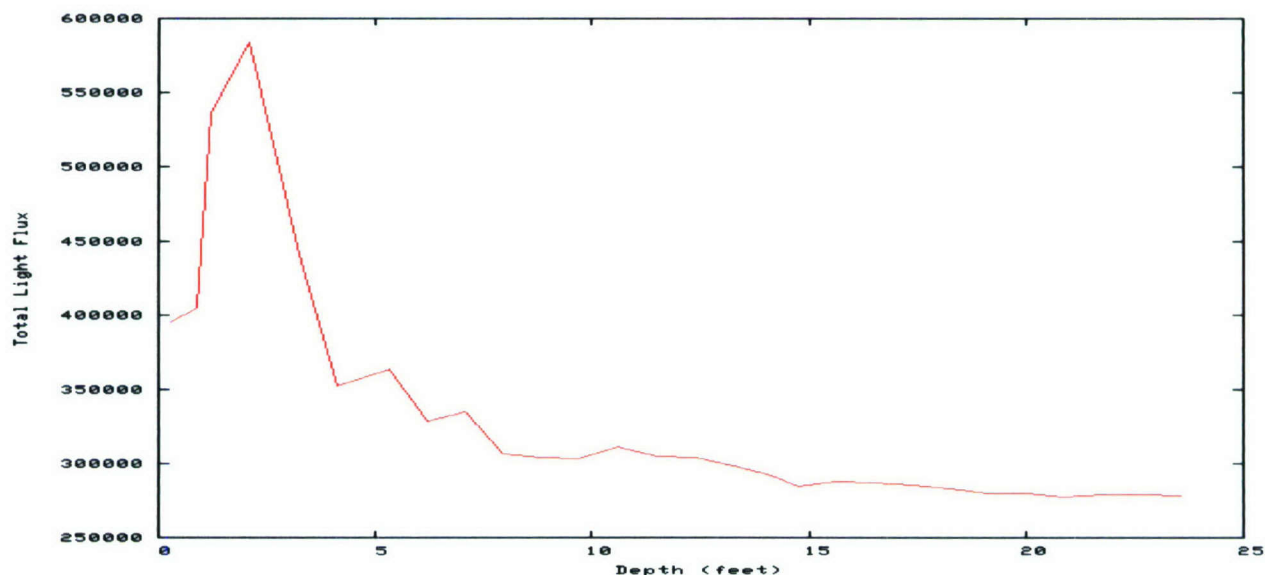
For this data, 700 nm (red) and 400 nm (blue) light are attenuated similarly, with an attenuation coefficient of 0.02/ft. Stated another way, 700 nm and 400 nm light are attenuated by 63% after traveling about 5 feet into the ocean. Green light, at 533 nm is attenuated by the smallest amount, with an attenuation coefficient of 0.1/ft (or a 10 foot penetration before reaching the 63% level).

The data we acquire here is unique in that we have data for so many wavelengths (200 to 1100 nm, in increments of 0.1 nm). We are working automated extraction and fitting software that computes the attenuation coefficients for *all* wavelengths, across the visible electromagnetic spectrum. A preliminary run resembles the following.



Here we have plotted attenuation coefficient versus wavelength. The data is unreliable for wavelengths less than 480 nm and greater than 650 nm for technical reasons which we are working to correct. It is our contention that the above plot is the "ultimate" in data reduction for this project: **a single graph that completely quantifies ocean color**. Here we see a minima in the attenuation coefficients at approximately 570 nm, indicating the ocean water most strongly sustains the propagation of 570 nm light. In other words, the ocean has the color of 570 nm light (or green). We plan on tracking the minima position, as well as its depth (contrast), and width over time.

**II. Light intensity with depth.** If we integrate under a given spectrum, such as the one shown at the top of page 2, we'd arrive at the number proportional to the total number of photons that existed at the given depth. If we then plot this total photon number versus depth, we arrive at a rather revealing plot of where light intensity seemed to exist at the time the data was taken. A plot taken in early November of 2006 resembles the following.



Of particular interest to us is why is the peak light intensity nearly 2.5 feet below the ocean surface? This plot clearly indicates an increase in light intensity with depth, leading to an eventual (and expected) exponential decay in light intensity. We will work to correlate the position of this peak with surface conditions to see if lensing of light due to waves is playing a role.

### **Presentation**

This work was presented at the December 2006 meeting of the American Geophysical Union (AGU) in San Francisco.

### **Conclusions**

We have developed and are currently operating an optical water column characterization tool, that collects sunlight samples at near continuous depths for a 30-foot water column. Data analysis is underway, that will quantify ocean color and a light intensity map with depth in the water column. We are grateful for the ONR and C3RP support for this work.

### **Publications and Presentations Related to Project**

#### **Abstracts**

*Bensky, T. J. and Neff, B.: 'Continuous optical monitoring of a near-shore sea-water column'; Presentation at American Geophysical Union, San Francisco, CA; December, 2006.*

Cal Poly San Luis Obispo runs the Central Coast Marine Sciences Center, south-facing, 1-km-long pier in San Luis Bay, on the west coast of California, midway between Los Angeles and San Francisco. The facility is secure and dedicated to marine science research. We have constructed an automated optical profiling system that collects sunlight samples, in half-foot increments, from a 30 foot vertical column of sea-water below the pier. Our implementation lowers a high quality, optically pure fiber cable into the water at 30 minute intervals. Light collected by the submersed fiber aperture is routed to the pier surface where it is spectrally analyzed using an Ocean Optics HR2000 spectrometer. The spectrometer instantly yields the spectrum of the light collected at a given depth. The "spectrum" here is light intensity as a function of wavelength between 200 and 1100 nm in increments of 0.1 nm. Each dive of the instrument takes approximately 80 seconds, lowers the fiber from the surface to a depth of 30 feet, and yields approximately 60 spectra, each one taken at a successively larger depth. A computer logs each spectra as a function of depth. From such data, we are able to extract total downward photon flux, quantify ocean color, and compute attenuation coefficients. The system is entirely autonomous, includes an integrated data-browser, and can be checked-on or even controlled over the Internet, using a web-browser. Linux runs the computer, data is logged directly to a MySQL database for easy extraction, and a PHP-script ties the system together. Current work involves studying light-energy deposition trends and effects of surface action on downward photon flux. This work has been funded by the Office of Naval Research (ONR) and the California Central Coast Research Park Initiative (C3RP).

Nonlinear Analytical Modeling and System Verification of Rigid Base and  
Seismically Isolated Wine Barrel Stacks

Project Investigators:

Charles. B. Chadwell

Assistant Professor

Department of Civil and Environmental Engineering  
California Polytechnic State University  
San Luis Obispo, California

D. Townsend Brown

Graduate Student Researcher

Department of Civil and Environmental Engineering  
California Polytechnic State University  
San Luis Obispo, California

# NONLINEAR ANALYTICAL MODELING AND SYSTEM VERIFICATION OF RIGID BASE AND SEISMICALLY ISOLATED WINE BARREL STACKS

C. B. Chadwell<sup>1</sup> and D.T. Brown<sup>2</sup>

## ABSTRACT

After poor performance of wine barrel stacks during the San Simeon Earthquake an initiative to find a safe and stable storage method became desirable. The current method of storing wine oak barrels stacked up to eight levels high on portable steel racks has been demonstrated during past earthquakes to have a high chance of collapse. When these wine barrel stacks are excited by seismic ground motion, they behave like a tall slender block susceptible to collapse. This intention of the research documented herein is both to experimentally and numerically evaluate wine barrel stacks from 2 to 5 barrels high using three selected ground motions representing a 10% chance of happening over a 50 year period of time. Experimentation with the different stack heights and ground motion records provide insight into the system response of the stack barrels. Analytical results suggest that for a 5 barrel stack, collapse will be protected provided a based friction coefficient of less than 13% can be achieved. Stacks of 4 and 5 barrels were tested on bearings providing an approximate coefficient of 10% and physical simulation showed the barrel stack to remain intact.

---

<sup>1</sup>Assistant Professor, Dept. of Civil and Environ. Engineering, California Polytechnic State University, San Luis Obispo, CA.

<sup>2</sup>Graduate Student Researcher, Dept. of Civil and Environ. Engineering, California Polytechnic State University, San Luis Obispo, CA.

## **Chapter 1**

### **Introduction**

---

This paper examines the non-linear behavior of portable wine barrel storage systems subjected to seismic ground excitations. In addition, the research, described herein explores a wine barrel stack hazard mitigation solution using roller bearings. The problem of rigid body rocking and sliding has been well studied and much of this previous research has been applied herein to wine barrel stack rocking/sliding. These rocking/sliding barrel stack response mechanisms are examined with both experimentation and a computer simulation model. This report documents the barrel stack testing, data reduction, test articulation with a nonlinear analysis model using typical structural engineering software, and stack collapse mitigation using a ball bearing seated in a Teflon surface reduce the force transference.

#### **1.1 California's Wine Industry**

California produces 450 million gallons of wine per year, which yields 45.4 billion dollars annually and over 200,000 jobs (Marrow, 2000). Numbers as large as this translate into a vested economic interest for many of the counties of California as well as the state. Wine requires extensive long term storage. The current standard for large scale storage is the portable two barrel rack. Warehouses are filled with barrel stacks, up to eight high, which are made by forklifts. Over 90 percent of all winery storage facilities are located in seismic zone 4 and a majority is within 20 miles of an active fault (Morrow, 2002).

The price of convenient storage has left the wine industry with a highly unstable storage method. Most wines age anywhere from six months to many years requiring a winery to possibly store large amounts of inventory at any given time. Given the historically poor performance of wine barrel stacks during seismic excitation, this poses serious risk to the economic stability of a winery. Past earthquake performance of wine barrel stacks have shown that these stacking devices are prone to instability and possible collapse during seismic ground shaking (Marrow, 2000). Unfortunately, awareness of the limitations of these sacking devices has only recently become a major concern (Morrow, 2000).

## **1.2 Previous Performance of Wine Barrel Stacks**

The 2003 San Simeon Earthquake caused significant damage to many of the local wineries both large and small (EERI, 2004). The damage was identified to be distributed heavily among the two barrel rack system (Figure 1). Nearly twenty local wineries in San Luis Obispo County reported collapsed wine barrel stacks, while no reports of structural damage to the storage facilities were reported. This indicates a discrepancy between current building code and heavy storage code.

## **1.3 Testing and Analysis of the Wine Barrel Portable Stacking System**

The objective of this research to gain an understanding of the seismic stability of wine barrel stacks through empirical data gained from full scale shake table testing with representative recorded ground motions and a comparison to a non-linear analytical model. The research conducted as part of the CCCRP (California Central Coast Research Partnership) encompassed testing of 2 barrel, 4 barrel and 5 barrel high wine barrel stacks. The stacks were instrumented at every barrel with position transducers and accelerometers to gain empirical data relating to the seismic performance of the wine barrel stack. This information is then compared to the output for the computer simulation model. Each test was videoed to assist in the data reduction procedure as well as simulation results.

In addition to the physical testing and model simulation, the research investigated a possible solution using steel ball bearings in a Teflon coated steel pad. The objective of this roller bearing is to reduce the friction



**Figure 1-1 Typical Wine Barrel Storage**



**Figure 1-2 Collapse of Wine Barrel Stacks**

transmission to the wine barrel stack thereby reducing the risk of stack collapse for a given ground excitation. The research described herein explains the analytical and testing iterations of the bearing developed.

## Chapter 2

### Literature Review

---

#### **2.1 Earthquake Rocking Response of Rigid Bodies**

Mohammad Aslam, A.M. ASCE, William G. Godden, M. ASCE and D. Theodore Scalise, Journal of the Structural Division, Feb 1980, pg. 377-392

In this study a computer model, BLOKROK, was created to simulate rigid body rocking. The program included boundary elements, aspect ratio, block size and the coefficient of restitution. Concrete blocks (6"x6"x30") were harmonically shaken and accelerations, velocities, rotations, and displacements were compared to the output from BLOKROK. In their analysis it was indicated that sliding allowed the block to be partly uncoupled with the horizontal component of ground motion. Furthermore, it was suggested that the base coefficient is the control quantity in these situations. They concluded that if the aspect ratio of the block is greater than the inverse of the coefficient of friction,  $1/\mu$ , then the block will not slide under the action of the ground motion. Conclusions of this study indicate the highly non-linear aspect of rigid body rocking. The rocking frequency and response is highly dependant on the boundary conditions. Furthermore rigid body rocking response is further dependant on aspect ratio, block size and coefficient of restitution.

#### **2.2 Rocking of Slender Rigid Bodies Allowed to Uplift**

Ioannis N. Psycharis, and Paul C. Jennings, Earthquake Engineering and Structural Dynamics, v.11, pg. 57-76, 1983

Examined in this paper is the rocking response of rigid blocks for the purpose of modeling the uplift that occurs in large buildings during seismic activity. The simple rigid block was modeled theoretically. The two models were used to establish the results: the Winkler foundation which uses an infinite number of springs along the base of the rigid structure and the simplified two spring foundation (springs located at the corner). Winkler's model and the two spring model use viscous dampers. The simplified two spring and two damper model placing dampers and springs at the outer pivot points were examined. Results of the study indicate that the two spring foundation is a simpler

approach to the problem, and can be used instead of the Winkler model. The actual numerical results from both methods are in fair agreement.

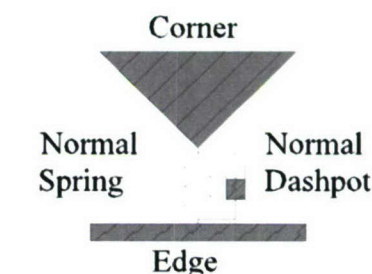
## **2.3 Response of Rigid Body Assemblies to Dynamic Excitation**

Tibor Winkler, Kimiro Meguro, and Fumio Yamazaki, *Earthquake Engineering and Structural Dynamics*, v.24, pg. 1389-1408, 1995

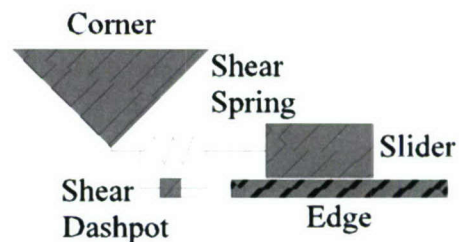
This paper compares a series of experiments to two analytical methods of rigid body rocking, one of which is more simplified Housner rocking block method. The experiment consisted of stacking rectangular wooden blocks in various arrangements and slenderness ratios, which were subjected to harmonic excitations. The mathematical model used in this research is a

combination of the work performed by Housner (1963) and a DEM (Distinct Element Method)

developed by Cundall (1971). A series of shake table experiments revealed various overturning accelerations and compared these results to two analytical methods (Figure 2-1, Figure 2-2). The DEM method incorporated more sliding and friction parameters that resulted in better agreement than the simplified Housner method. The distinct element method was analyzed and showed good agreement with the non-linear behavior of the rigid body rocking tests at low frequencies. Results did not have agreement at frequencies above 3 Hz.



**Figure 2-1 Damping element of the DEM Rocking Model**



**Figure 2-2 Friction Slider element of the DEM Rocking Model**

## **2.4 Seismic Response of Equipment Anchored to a Base Foundation Allowing Uplift**

Makris, N., C.J. Black, (2001) "Rocking Response of Equipment Anchored to a Base Foundation". PEER 2001/14, Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Outlined in this paper is the formulation of the rocking response of stored rigid bodies anchored to a foundation. A typical concern is that large devices such as electrical equipment have the ability to rock and possibly topple over during earthquakes. The main

focus of the paper is the minimum anchor strength required to support these objects from overturning. Both the service level and ultimate design level earthquakes are examined. Slenderness effects are considered by looking carefully at the plan dimensions verse the footprint of the equipment. Conclusions to the study indicate that there are two distinct capacities to resist uplift: the anchorage must be strong enough to engage the foundation into rocking and if the restrains fail the object will rock as a free standing block.

## ***2.5 The Rocking Spectrum and the Limitations of Practical Design Methodologies***

Nicos Makris and Dimitriou Konstantinidis, Earthquake Engineering and Structural Dynamics 2003, v.32, pg. 265-289

This paper outlines the fundamental differences between the single degree of freedom (SDOF) oscillator and the rocking response of a rigid block. An attempt is made to prove that the present FEMA 356: Prestandard and Commentary for the seismic rehabilitation of Buildings should abandon the equivalent SDOF oscillator method of determining the response of rocking blocks. The systematic approach outlines the fundamental similarities as well as the differences in the response of the two systems. While the response spectrum has become a mainstream tool in earthquake engineering, Makris claims that it is insufficient for rigid body rocking, especially with low slenderness ratios. In combination with the rocking spectrum, “a more lucid picture of kinematics characteristics” can come to light. Assumptions in the paper require an energy loss due to pounding as the block rocks back and forth. The Paper builds on previous works from Housner (1963) and his work on, “The Behavior of Inverted Pendulum Structures during Earthquakes.” Further assumptions include; the minimum energy loss during impact depends on the slenderness ratio of the block.

## ***2.6 Nonlinear Analytical Modeling and System Verification Utilizing a Combined Slide-Rock Seismic Response of Wine Barrel Stacks***

C. B. Chadwell, J.M. Stanley and J. Marrow, Earthquake Research Institute, Proceedings of the 13<sup>th</sup> World Conference on Earthquake Engineering, Vancouver, B.C., Canada. August 2006.

This research, which was performed at Cal Poly, SLO, in 2005, investigated the single wine barrel (on a steel rack) seismic behavior through full scale testing and a nonlinear time history analytical program written in Matlab. Sliding and rocking behaviors were observed in the shake table test and then modeled. The ground motions in the full scale test included a series of sine sweeps and the three SAC ground motions (LA16, LA18, and LA19). Fair correlation was achieved between the experimental tests and the

analytical model. System damping proved to have a minimal effect on the rocking behavior. The friction coefficient between the ground and rack had a controlling effect on the system behavior. Two modes of vibrations were identified for the on barrel on rack. The first, 0.05 sec, was independent of ground motion (for 4Hz, 6Hz, and 8Hz sine sweep ground motions) and correlated to the rocking period. The second period, 0.35 sec, was recognized in the 4Hz sine sweep test. The analytical model included material properties of the steel rack, rigid body wine barrel with a lumped mass at the centroid, Raleigh viscous damping, and gap/friction elements at the bottom corners of the rack (points of rocking).

## **2.7 Experimental Studies on the Earthquake Performance of Wine Barrel Stacks**

Marrow, J.M., Earthquake Research Institute, Proceedings of the 7<sup>th</sup> National Conference on Earthquake Engineering, Boston, Mass, July 2002.

This paper discusses the wine industry in California and the current storage methods for wine. Through a series of half scale and full scale wine barrel shake test he concludes the seismic instability of the current stacking methodology. His research also included destructive testing of wine barrels due to free falls and shear capacity of a wine barrel rack. Wine barrel leaked slightly from a 4.04ft fall (2<sup>nd</sup> barrel height) and all testing above this height resulted in complete destruction of the barrel. He distinguishes six failure mechanisms for a wine barrel stack subjected to ground excitation (under six degrees of freedom). Five of the six mechanisms occur from ground motion in the longitudinal axis direction (horizontal movement from barrel head to barrel head). Marrow concludes that the rocking response and sliding of the entire rack are the desired mechanisms. Rocking dissipates the most energy and slide mitigates the forces transmitted to the stack.

## Chapter 3

### Experimental Procedure

---

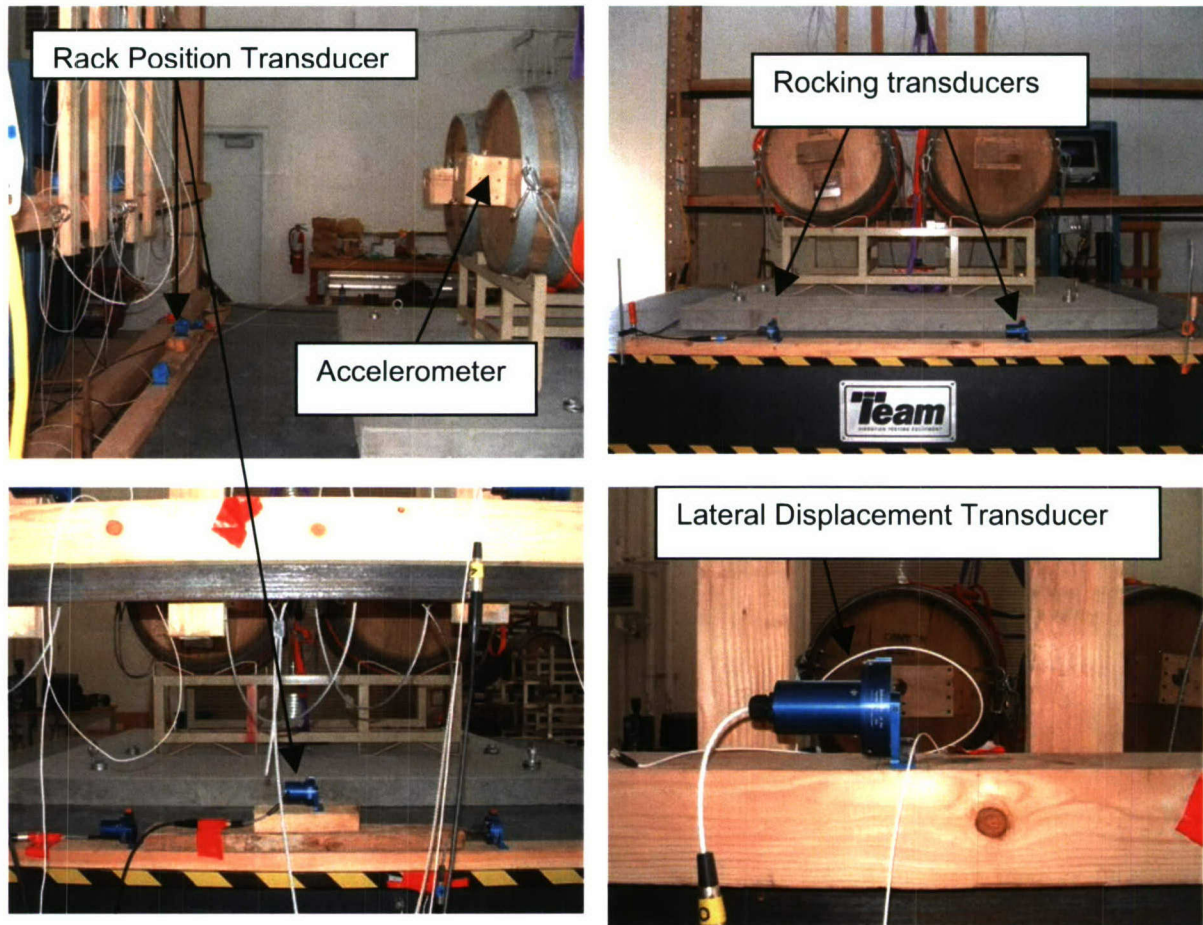
The purpose of the full scale shake table testing is to gain insight into the rocking and sliding mechanisms of wine barrel storage system during earthquakes. The test setup includes the typical steel rack and oak barrel stacking method, data acquisition instrumentation, a safety restraint system, and an equipment frame for mounting external measuring devices. A concrete slab was designed and installed on the shake table to imitate the typical cellar floor during year 1 of the research and testing. Position transducers and accelerometers were attached to the test specimens in order to collect data later used for model calibration. The barrel-rack stacks were then placed on the concrete surface completing a typical storage scenario, and earthquake time histories were conducted on the shake table.

#### **3.1 The Testing Facility and Data Acquisition Devices**

All experimental testing was performed at the Parsons Earthquake and Geotechnical Engineering Laboratory in the Advanced Technologies Laboratory at California Polytechnic State University, San Luis Obispo. This is the location of the 10 ft by 10 ft Team unidirectional shake table. The table has a capacity of +/- 5.25 inches, 38 in/s<sup>2</sup> maximum velocity and a peak acceleration of 10g. The table operates from a central computer, frequency generator, and hydraulic actuator. Various earthquake time histories can be programmed and run though the table, provided that they remain within the limits of the table capacity.

Data collection for the experiment utilized a National Instruments Data Acquisition system. Data acquisition included lateral displacement measurements, vertical displacement measurements between barrels and lateral acceleration. Each device operates on an individual channel within the operating software. Due to the nature of the test protocol, a high frequency collection frequency was utilized (400 Hz). Displacements were measured with Spaceage Control position transducers. Three types of transducers were used with various ranges and accuracies. The shorter range, more sensitive, transducer type was used to measure displacements between wine barrels in the vertical direction, where smaller a displacements range was needed. Accelerometers were mounted on each barrel in the barrel stack tests. The accelerometer measured accelerations up to 20g with a 0.01g accuracy in a 2D plane. The reasons and details on

the location of the instrumentation are more thoroughly covered in the data reduction section.



**Figure 3-1 Pictures of Laboratory Instrumentation**

### **3.2 The Wine Barrel and Rack**

Three different size barrels were donated for this project. The dimensions are given in Figure 3.2. French Bordeaux barrels were filled with water, which has a negligible difference in specific gravity from wine. These barrels are constructed from oak staves shaped into a bulging cylinder with hoops around it, a flat circular head at both ends, and at least one hole, generally on the belly, for a bung (barrel plug). Most barrels weigh between 65 lb. and 100 lb. empty and approximately 575 lb. full of wine. 6 to 8 steel hoops are generally spaced along the barrel length. Wineries can specify the placement of

the end hoops to either be at the contact point of the rack cradles or to be on either side to allow for a bearing point of metal rack to oak barrel (Morrow 2002).

The steel racks used for the experiment were the two barrel industry standard. Topco model WR2 steel racks were furnished for experimental testing. The WR2 rack is constructed of ASTM A36 steel 1 1/2" steel box sections (0.065" thick). The saddle type used for testing was the "Eastern

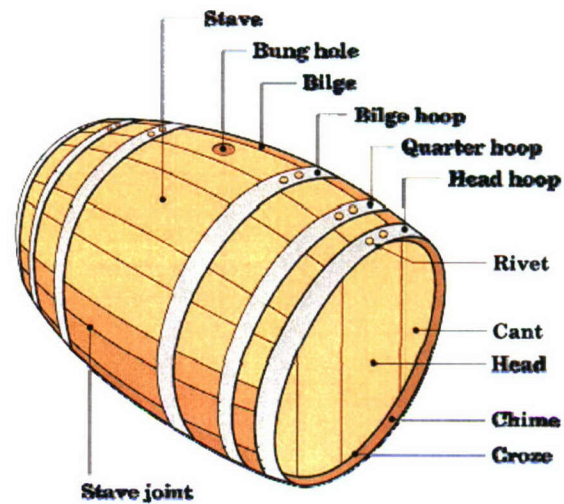
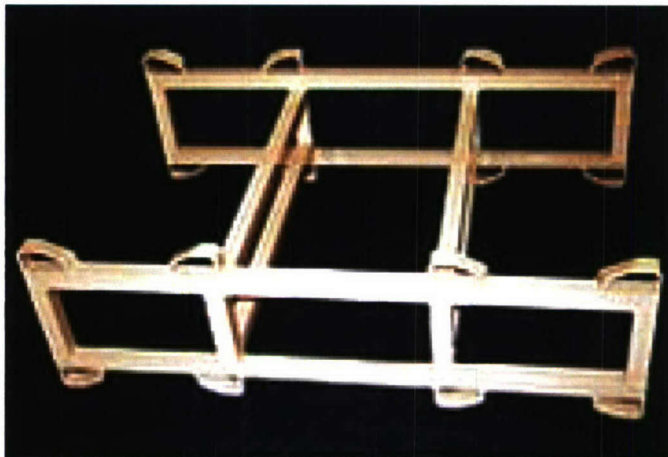


Figure 3-2 Barrel Parts (www.rootsweb.com)

Table 3-1 Barrel Dimensions

Barrel Volume	Properties	Dimensions (inches)	Dimensions (mm)
59 Gallon/ 225 Liter	Stave Thickness	0.987"	24 mm
	Diameter of Head	22.687"	58 mm
	Circumference of Head	71.27"	181 mm
	Diameter of Belly	27.53"	70 mm
	Circumference of Belly	87"	221 mm
	Barrel Height	34.875"	89 mm
	Diameter of Bung Hole	2.0"	50 mm
53 Gallon/ 200 Liter	Stave Thickness	0.937"	24 mm
	Diameter of Head	21.25"	54 mm
	Circumference of Head	66.8"	170 mm
	Diameter of Belly	25.7"	65 mm
	Circumference of Belly	80.625"	205 mm
	Barrel Height	34.875"	89 mm
	Diameter of Bung Hole	2.0"	50 mm
65 Gallon/ 246 Liter	Stave Thickness	0.937"	24 mm
	Diameter of Head	22.687"	58 mm
	Circumference of Head	71.27"	181 mm
	Diameter of Belly	28"	71 mm
	Circumference of Belly	88"	224 mm
	Barrel Height	34.875"	89 mm
	Diameter of Bung Hole	2.0"	50 mm

Saddle” and is used most often in industry applications (Figure 3.3). The racks are powder coated by the manufacture to provide superior corrosion resistance and aesthetic appeal. The overall dimensions of the racks are 30-1/2” in length, 44-1/2” in width, and have a 7” opening for fork lift clearance. They can support two wine barrels and stack heights of at least six to eight barrels high under gravity static loading. Since wine barrels are offered in various sizes depending on application, these racks can be used with either of the barrel sizes; however, mismatching of barrels on the same level in a stack results instability of the stack and improper seating.



**Figure 3-3 Topco WR2**

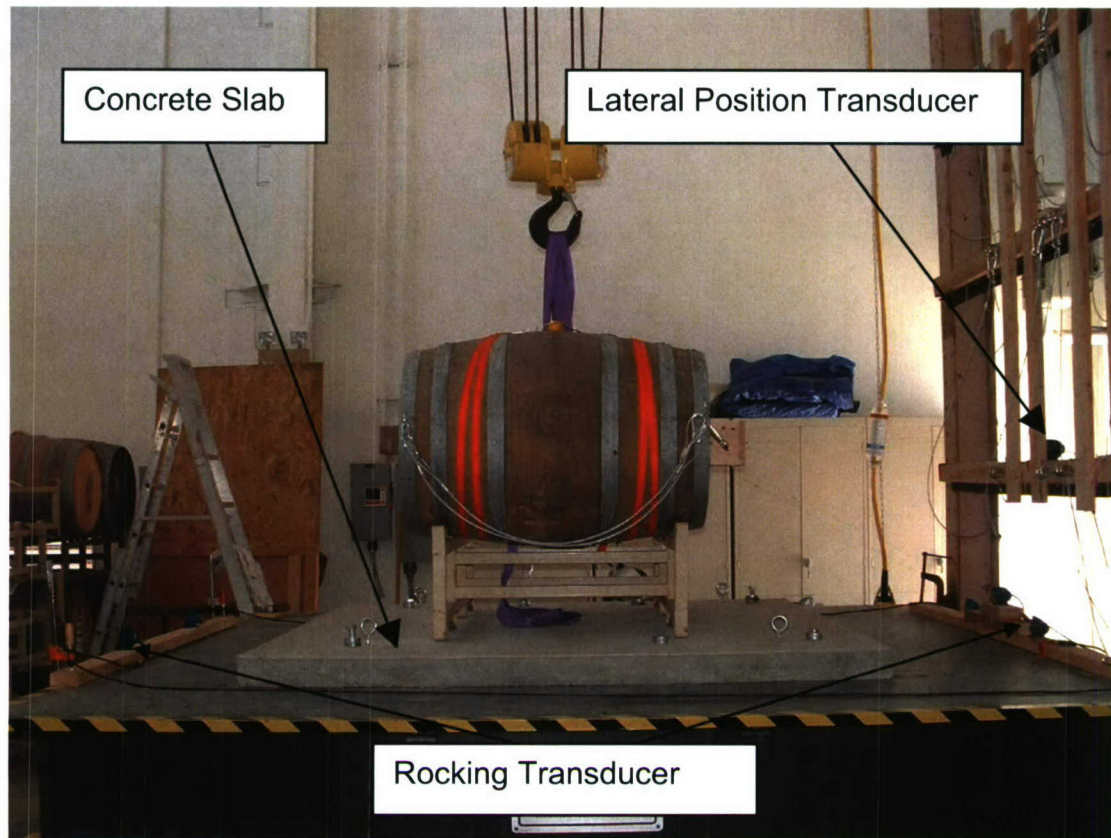
### **3.3 Existing Experimental**

#### **Setup**

The one barrel stack tests set the framework (Figure 3-4) for the two, four and five barrel stack testing as part of the CCCRP funded work in 2005. During the year 1 phase of this project:

- An instrument frame was built enabling vertically fixed locations for the horizontal position transducers
- A four inch concrete slab was mounted on the shake table to mimic the average slab in a wine cellar.
- The shake table was calibrated with the slab as the dead load.
- Instrument boxes were constructed to protect the vertical position transducers and the accelerometers, which were mounted on the right barrels and the bottom left barrel (centered on the head of the barrel).
- A redundant safety system designed and installed to prevent the barrels from free fall, but enable unhindered movement of the barrels until collapse. This consisted of steel cables and carabineers connecting barrel to barrel and barrel to instrument frame.

With previous laboratory set up completed and a consistent barrel stacking procedure, everything was in place for a smooth transition to higher stacking.



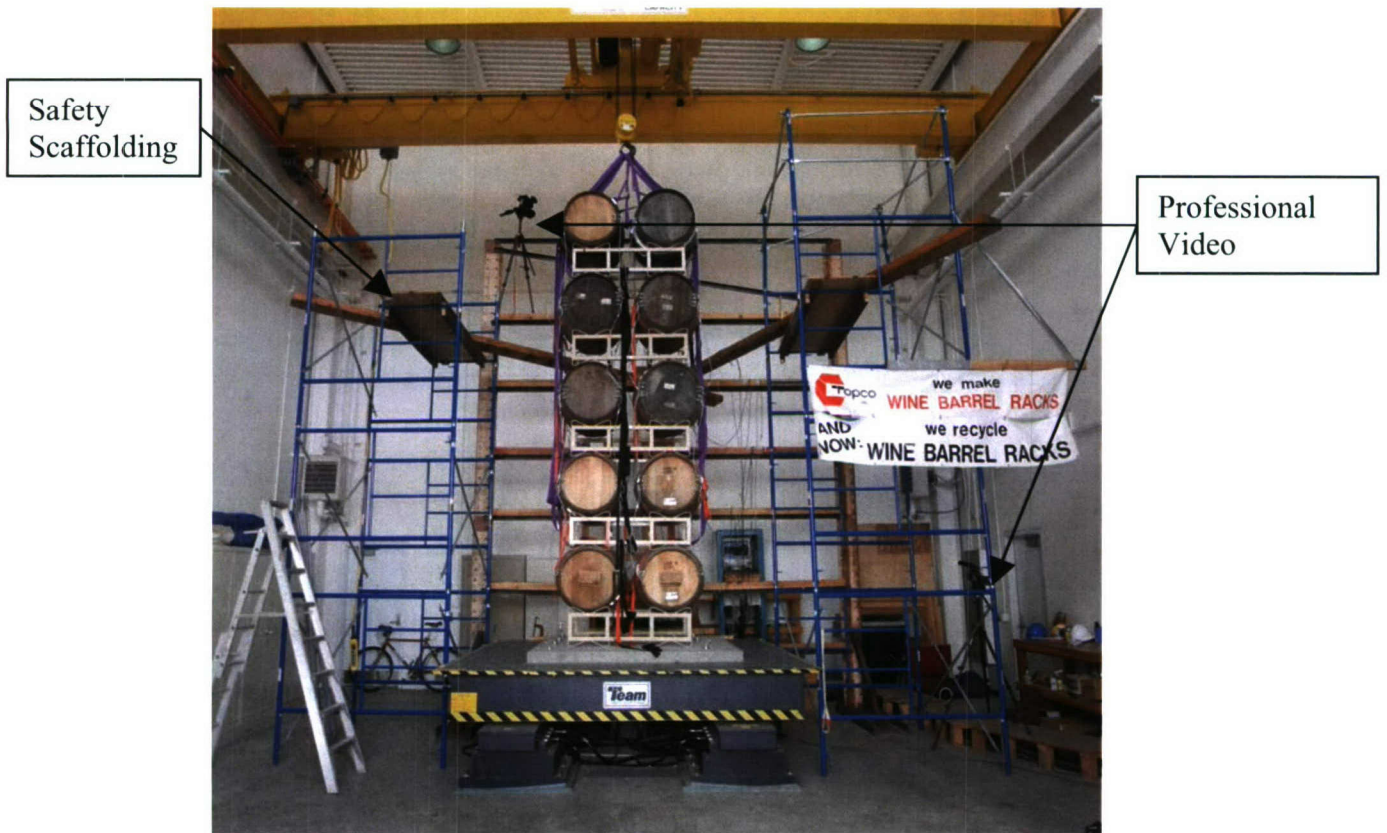
**Figure 3-4 One Barrel Set-up**

### ***3.4 Additional Setup for the Five Barrel Stack***

Two scaffolding towers were constructed to facilitate five barrel high stacking and two 4" by 10" wood beams were used to temporarily connect the scaffolding towers. To aid in the data reduction process professional video was taken at each test from two angles (Figure 3-5). These angles were synchronized with each other when the DVD was edited.

### ***3.5 Testing Procedure***

For each stack height (2 barrel, 4 barrel, and 5 barrel) three different ground motions were tested, the LA16, LA18, and LA19. The output was record in a text file and then reduced to obtain barrel translation and rotational data.



**Figure 3-5 Present Experimental Setup**

### **3.6 Earthquake Records for Tests**

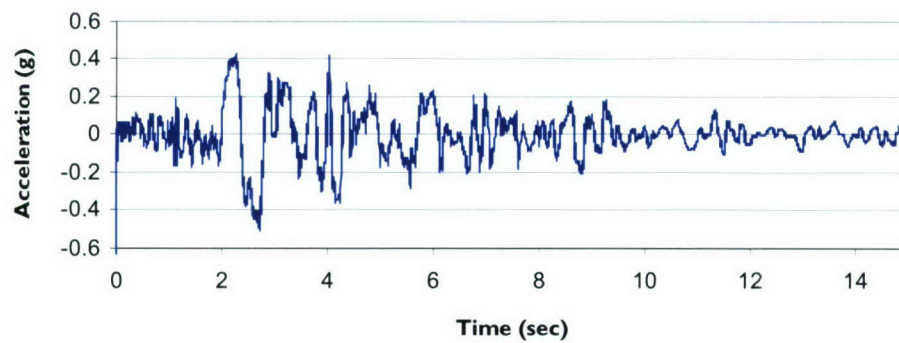
Three earthquake ground motions were tested on the barrel stacks to investigate the tendency for nonlinear rocking/sliding model by capturing the measured uplifts and displacements. The ground motions were originally chosen (Chadwell et, 2006) for their individual characteristics and frequency content to be consistent with previous tests, the same three ground motions were used in the current experiments. The ground motion records containing shorter period energy were selected to represent rocking type behavior while the ground motion records containing longer period energy were chosen in an effort to induce sliding in the model. The records selected are classified as earthquakes that have a 10% probability of occurring over a 50 year period (SAC, 1996). Details of the ground motions selected (LA16, LA18 and LA19) and their respective acceleration records are provided in Table 3.2 and Figures 3.6-3.8.

**Table 3.2 Ground Motion Characteristics****Details of Los Angeles Ground Motions Having a  
Probability of Exceedence of 10% in 50 Years**

Name	Record	Earthquake Magnitude	Dist. (km)	Scale Factor	Number of Points	DT (sec)	Duration (sec)	PGA (cm/sec <sup>2</sup> )
LA16	Northridge, 1994, Rinaldi RS	6.7	7.5	0.79	2990	0.005	14.945	568.58
LA18	Northridge, 1994, Sylmar	6.7	6.4	0.99	3000	0.02	59.98	801.44
LA19	North Palm Springs, 1986	6	6.7	2.97	3000	0.02	59.98	999.43

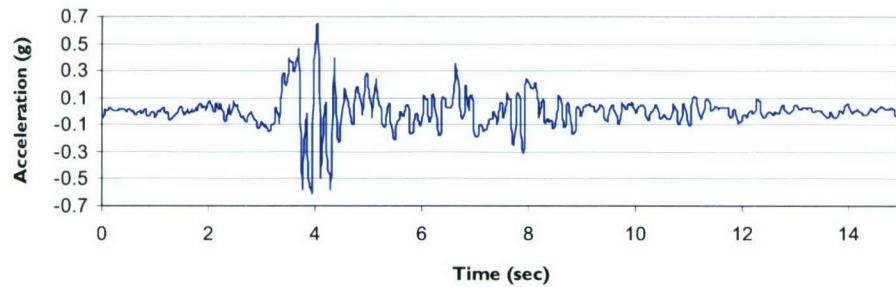
Figure 3.9 shows the linear response spectrum calculated at 5% of critical damping for the three ground motion records selected. Barrel tests (and analysis) indicate the system (two, three, and four barrel stacks) has an elastic fundamental period of less than 1.0 sec. At this period, all three ground motions are in the acceleration controlled region of the spectrum. However, LA16 has a significantly lower acceleration at this period than do the LA18 and LA19 ground motion records.

**LA16 Acceleration Time History**



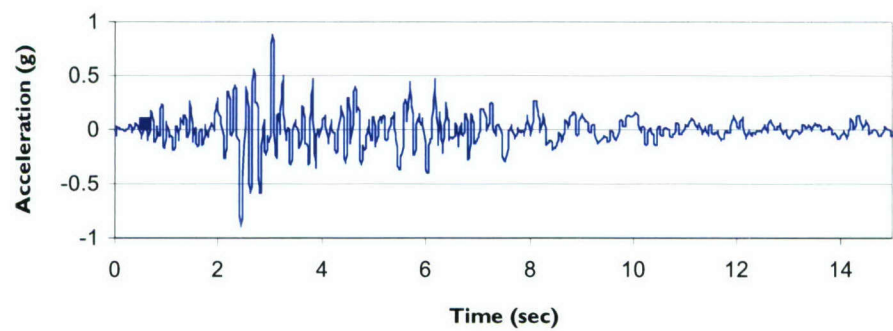
**Figure 3.6 - LA16 Acceleration Time History**

**LA18 Acceleration Time History**

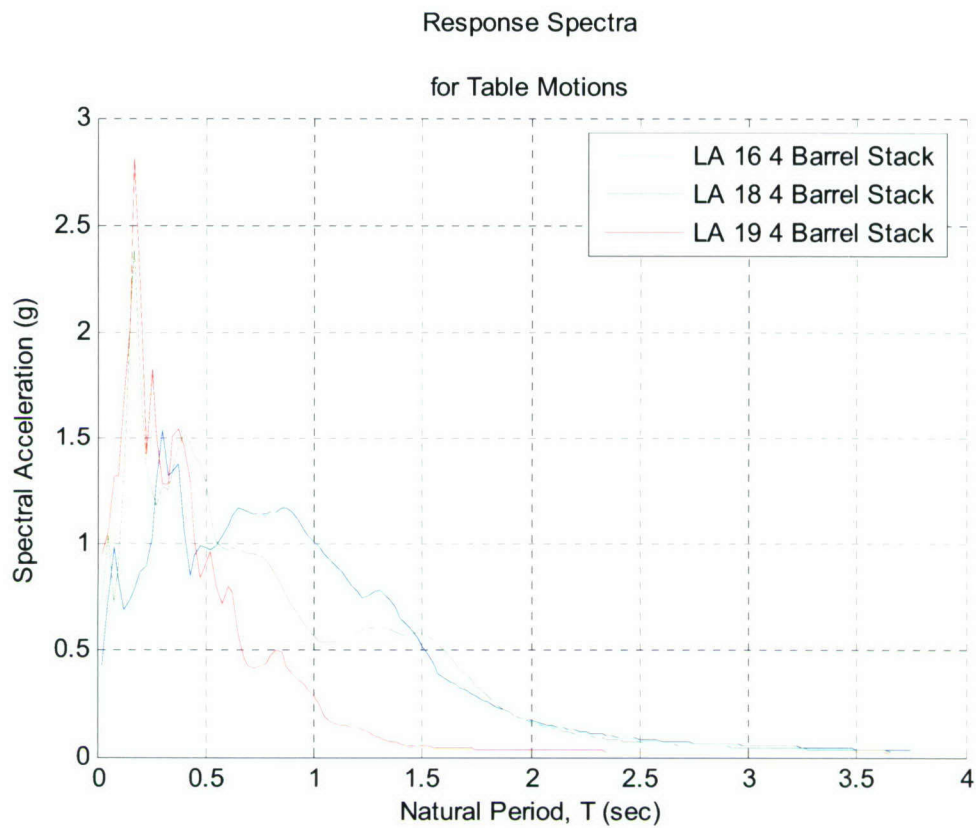


**Figure 3.7 - LA18 Acceleration Time History**

**LA19 Acceleration Time History**



**Figure 3.8 - LA19 Acceleration Time History**



**Figure 3.9 Response Spectra for 4 Barrel table motions**

## Chapter 4

### Data Reduction

---

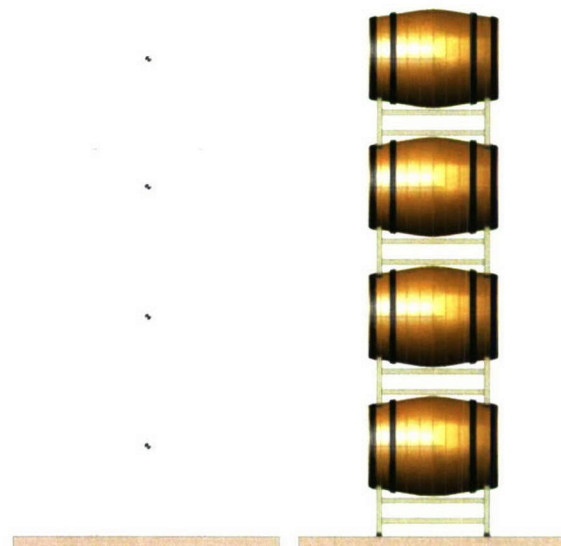
From each shake table test, the data acquisition system recorded information from 13 position transducers and 6 accelerometers at 400 hertz. The responses of interest for the data reduction process are rocking and sliding of each barrel, accelerations of each level and rotation of the stack. This was gained through position transducers placed on a frame horizontal to each barrel and vertically between barrels. This section is broken into three parts; the first will explain the physical steps made in reducing the information, the second will address the theory, assumptions and its implications in reducing the data and the third will explicitly describe the formulas used.

#### **4.1 The Data Reduction Process**

The data acquisition system's output included 20 columns with 15,000 to 55,000 rows of data. This array included time, table position, 7 horizontal transducers, 6 vertical transducers, and 5 accelerometers all writing at 400 hertz. Non critical pre and post earthquake test data was removed from the data files and the transducer and accelerometer readings were zeroed and labeled. The data was manipulated to find the vertical, horizontal, and rotational movement of the barrel's centroid each barrel at every time step.

#### **4.2 The Assumptions and Their Implications**

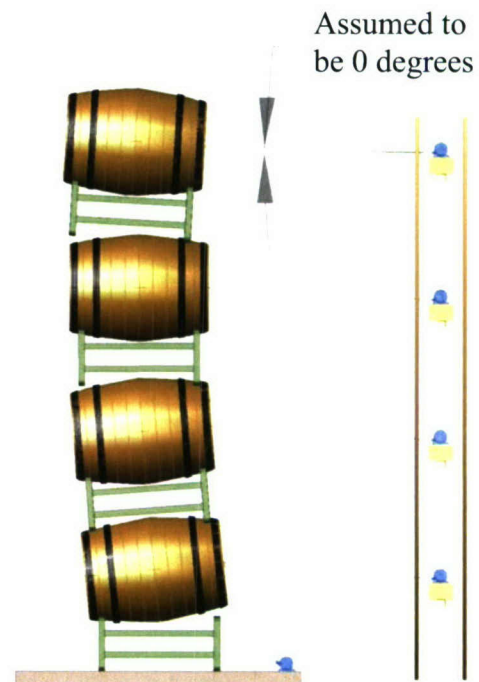
The interface between the wine barrel and the rack below it was assumed to be rigidly connected. This assumption was confirmed through analysis of test video where each barrel to rack interface was inspected throughout each experiment. There was never more than approximately 0.25 inch of movement a wine barrel and its adjacent supporting rack. This movement was small compared to the movement between the rack and the barrel below. This assumption however, did not apply to the top barrel and



**Figure 4-1 Data Reduction Model**

the one barrel high stack test. All rocking was assumed to occur at the bottom corners of the rack. This held true until complete uplift of a barrel-rack from the stack below. These cases were limited to the LA 16. This data reduction model simplifies the problem dramatically, because there are only be five bodies to identify instead of ten in a five high stack. The rack is assumed to be centered under the barrel above it. An error analysis would not be appropriate because no horizontal or vertical transducers were attached to the racks (excluding the bottom rack).

Because of the symmetry, the racking system, and one degree of freedom shaking, it is assumed that all barrels at the same level must rotate and translate in unison as one module. Only one of the two columns of barrels was instrumented requiring this assumption to be made.



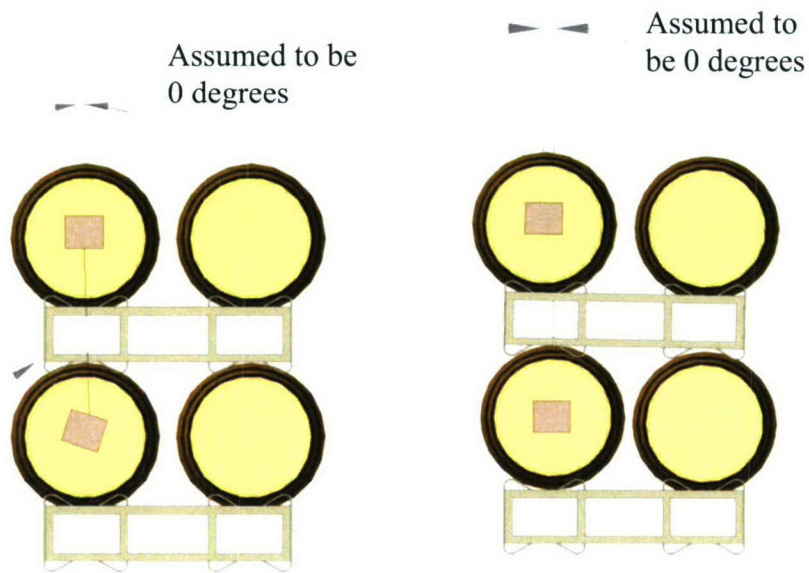
**Figure 4-2 Rotation Assumption 3**

The horizontal transducers were assumed to stay perfectly horizontal. The horizontal distance between the barrels and the instrumentation frame varied between 8-9 ft and the horizontal transducers were attached to each barrel ( $\pm \frac{1}{2}$  inches vertically). Throughout each test, the barrels would not move vertically more than 5.5 inches thus resulting in a minimal error in this assumption. This was deemed acceptable up until collapse of the barrels, because the horizontal transducers never deviated more than 3 degrees from true horizontal giving an acceptable maximum error in the horizontal transducer of 0.15 inches.

Barrel rotation in the transverse direction from table movement was ignored. Only in a few tests was rotation noticed in this direction. This rotation would induce errors in the vertical transducer measurements (Figure 4-3). However, the vertical transducers never deviated more than 2 degrees from vertical resulting in an acceptable maximum error of 0.018 inches.

Due to warping in the racks and inconsistencies in the barrel shapes, the barrels did not rest in plane vertically (in elevation view, Figure 4-4). This creates error in the vertical transducers. Due to this source of error, the vertical transducers never deviated more than 1.5 inches horizontally resulting in a maximum error of 0.035 inches.

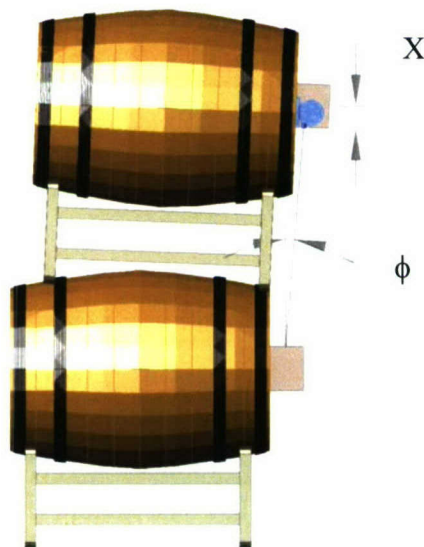
The instrumentation box was assumed to be centered on the barrel with both vertical and horizontal transducer wires attached to that point from barrel to barrel and barrel to instrument frame, respectively. Because of the finite transducer size, the top barrel's wire



**Figure 4-3 Barrel Assumption 4**

**Figure 4-4 Barrel Assumption 5**

does connect to the center. This would not create any error if the barrels stayed in plane profile, but it does create an error when the barrels are offset during ground shaking (Figure 4.5). This error was also deemed to be insignificant. The maximum error (if all barrels were of equal size) was found to be 0.008 inches. This error is magnified with barrels of different lengths (+1.25 inches horizontally) resulting in a maximum error of 0.12 inches.



**Figure 4-5 Transducer Box Error**

A summary of the assumptions and corresponding errors are listed in Table 4-1.

Table 4-1 Assumption Recapitulation

Assumption #	Instrument Output in Question	Range in Error
1	None	N.A.
2	None	N.A.
3	Horizontal Transducer	+0.15"
4	Vertical Transducer	+0.018"
5	Vertical Transducer	+0.035"
6	Vertical Transducer	+0.125"

## 4.3 The Explicit Relationships

### 4.3.1 Definitions

For all the data reduction described herein, the variables used are defined as:

- $h_b$  is the diameter of the barrel at the location where the barrel rests on the rack.
- $h_r$  is the vertical distance between the ground and the rack to barrel contact.
- $w_r$  is the width of the rack measured from outside corners (rotation points).
- $a(t)$  is the horizontal distance from the corner of the rack to the position transducer wire at any time  $t$ .

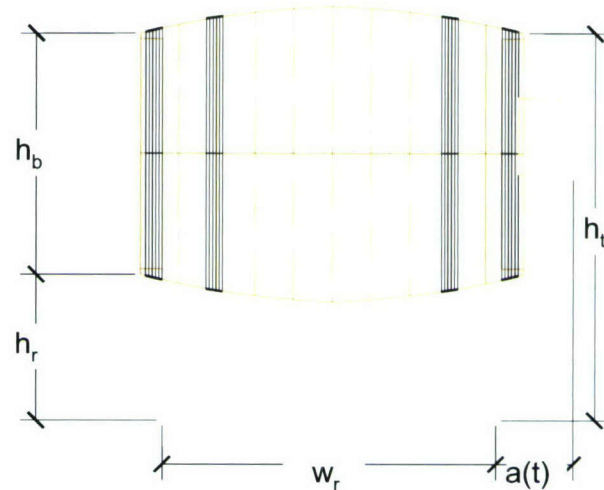


Figure 4-6 Barrel Dimensions

The height of the barrel, the vertical distance between the ground and the barrel contact point, and the width of the rack are stationary in time and are assumed the same for every barrel for every test.

At each measured time  $t$  during the test, the horizontal movement of one barrel over the next is a sum of barrel sliding plus barrel rocking. Thus the difference in recorded transducer data between barrel heights can be defined in terms of a sliding and rocking

displacement. Defining  $\Delta x_{\text{trans}}$  as the total difference in horizontal displacement between barrel centroids, it is the sum of horizontal motion due to rocking at the instrument box,  $\Delta x_{\text{rock\_Box}}$ , and horizontal motion due to sliding measured from the instrument box,  $\Delta x_{\text{slide\_Box}}$ .

$$\Delta x_{\text{trans}} = \Delta x_{\text{rock\_Box}} + \Delta x_{\text{slide\_Box}}$$

Assuming small angles, the horizontal sliding of the box is equal to the horizontal movement of the centroid of the barrel,  $\Delta x_{\text{slide}}$ . The angle created by  $\Delta x_{\text{slide}}$  is due to the shape of the barrel and never exceeds 3 degrees. Thus,

$$\Delta x_{\text{slide\_Box}} = \Delta x_{\text{slide}}$$

When the barrel rocks the instrumentation box and the centroid of the barrel are no longer horizontally in line with each other, causing a horizontal displacement on the barrel,  $\Delta x_{\text{rock}}$ . The magnitude of horizontal displacement due to rocking is dependent on the distance from the centroid to the point of rotation (bottom corner of the rack). When the barrel rotates clockwise the distance between the transducer connection and the rotation point (the corner of the rack) is roughly the same ( $\pm 1.0$  inches) as the distance from the barrel's centroid to the corner of the rack. With counter-clockwise rotation the horizontal distance from the rotation point to transducer connection is over twice the distance to the centroid. Thus, for clockwise rotation:

$$\theta_{\text{rock}} < 0: \Delta x_{\text{rock\_Box}} = \Delta x_{\text{rock}}$$

While for counter clockwise rotation:

$$\theta_{\text{rock}} > 0: \Delta x_{\text{rock\_Box}} = \cos(\theta_{\text{rock}}) * \Delta x_{\text{rock}}$$

The vertical transducer's output has been broken-up in a similar fashion with one additional term. Because the vertical transducers are connected to the individual barrels and not a fixed frame, the transducer can move in two dimensions and an adjustment to the output needs to be made. This adjustment is named  $y_{\text{error}}$ . Before any barrel performance is calculated, the error in the vertical transducer at each barrel number (#),  $y_{\text{error}\#}(t)$ , must be taken out at each time step. The total change in displacement in the transducers vertically connected between the wine barrels is the sum of the error, the horizontal barrels sliding measured at the box level, and barrel rocking (also measured at the box level). Thus the transducer reading can be taken as:

$$\Delta y_{\text{trans reading}} = \Delta y_{\text{slide\_Box}} + \Delta y_{\text{rock\_Box}} + y_{\text{error}}$$

Where  $y_{\text{error}}$  (dependent only on the barrel below) can be taken as:

$$y_{\text{error}} = \sqrt{d_{\text{unit}}^2 + \Delta x_{\text{slide}}(t)^2} - d_{\text{unit}} - \Delta y_{\text{slide}}(t)$$

Where  $d_{\text{unit}}$  is the original vertical distance from instrument box to instrument box. A summary of the variable definitions is shown in a schematic form in Figure 4.7.

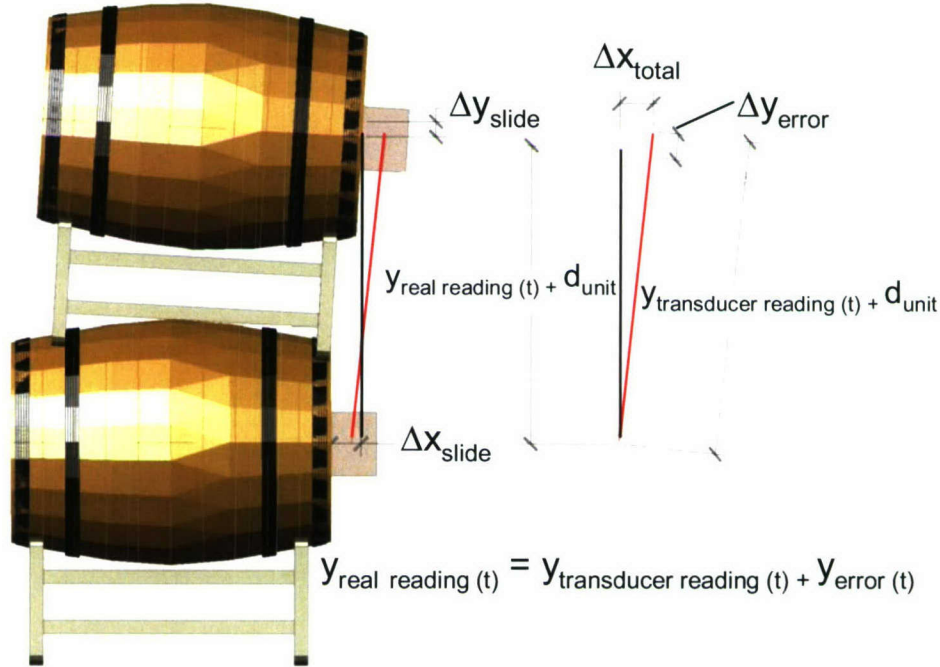


Figure 4-7 Variable Definitions for Calculating Barrel Translations

### 4.3.2 Barrel Rotation

Rotation of a wine barrel due to seismic excitation can occur from rocking and sliding. Rocking obviously contributes to a rotation of the barrel, but the barrel geometry is such that sliding also contributes to barrel rotation. When the barrel-rack slides along the curvature of the wine barrel below, a rotation is induced. Defining the total individual  $i^{\text{th}}$  barrel rotation as the sum from rotation due to sliding ( $\theta_{\text{slide}}$ ) and rotation due to rocking ( $\theta_{\text{rot}}$ ):

$$\theta_i = \theta_{\text{slide}} + \theta_{\text{rot}}$$

Then, the total barrel rotation,  $\theta_{\text{total } i}$ , is the total barrel rotation of the  $i^{\text{th}}$  barrel as defined by the sum of individual total rotation contributions starting at the  $i^{\text{th}}$  barrel and including all the barrels below (Figure 4-8). For example:

$$\theta_{\text{total } 2} = (\theta_{\text{slide } 1} + \theta_{\text{rot } 1}) + (\theta_{\text{slide } 2} + \theta_{\text{rot } 2})$$

$$\theta_{\text{total } 3} = (\theta_{\text{slide } 1} + \theta_{\text{rot } 1}) + (\theta_{\text{slide } 2} + \theta_{\text{rot } 2}) + (\theta_{\text{slide } 3} + \theta_{\text{rot } 3})$$

$$\theta_{\text{total } 4} = (\theta_{\text{slide } 1} + \theta_{\text{rot } 1}) + (\theta_{\text{slide } 2} + \theta_{\text{rot } 2}) + (\theta_{\text{slide } 3} + \theta_{\text{rot } 3}) + (\theta_{\text{slide } 4} + \theta_{\text{rot } 4})$$

The rotation is found from a contribution of the vertical transducer reading. Because the vertical position transducer is only on one side of the barrel, rotation orientation is dependent on positive or negative output readings. Thus,

$$\Delta y_{\text{trans rock}} > 0 \text{ then } \theta_i = \text{CCW or positive}$$

$$\Delta y_{\text{trans rock}} < 0 \text{ then } \theta_i = \text{CW or negative}$$

Defining  $a(t)$  as the horizontal distance, referenced from the start of the experiment (time zero), from the corner of the rack (the rotation point) to the transducer location.

$$a(t) = a(0) - \Delta x_{\text{slide}}(t)$$

The rocking rotation can be found as a function of either CW or CCW barrel rocking. Figures 4-9 and 4-10 schematically depict the variables for calculation of CW and CCW rotation.

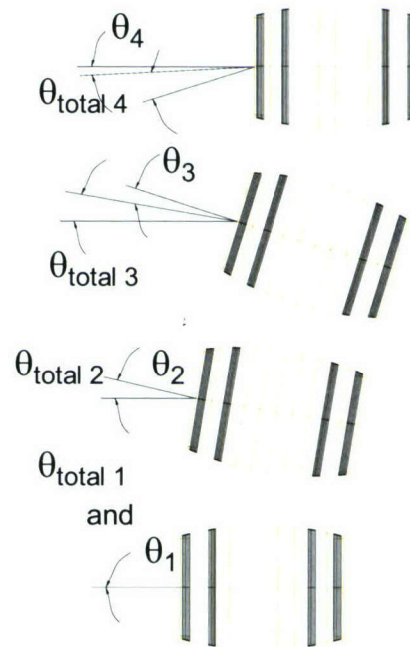


Figure 4-8 Rotation Nomenclature

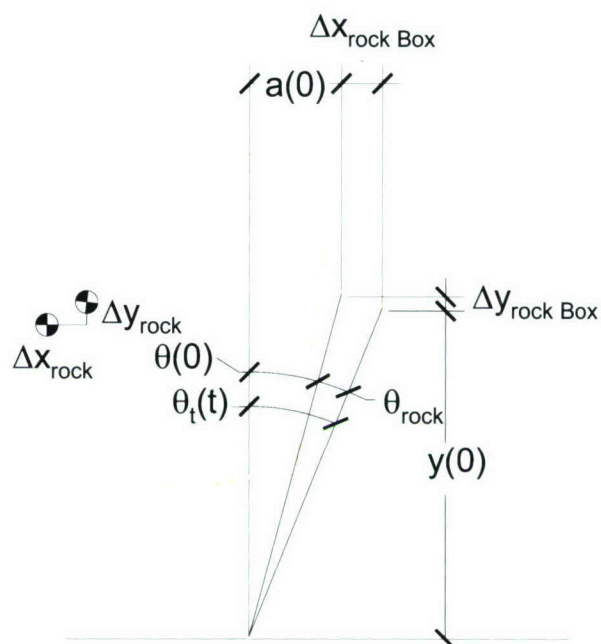


Figure 4-9 CW Rotation

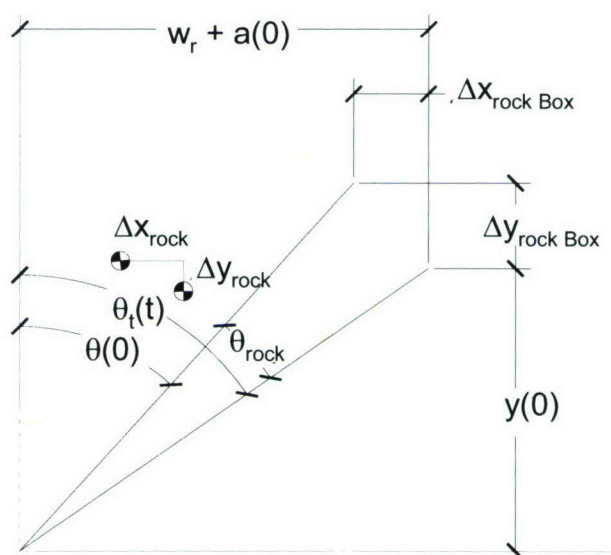
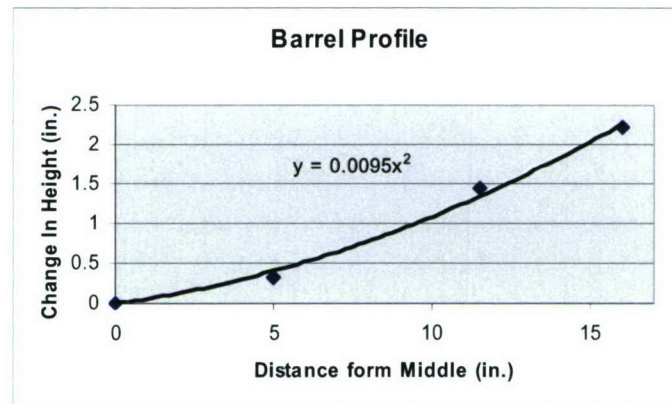


Figure 4-10 CCW Rotation

For the rotation due to sliding,  $\theta_{\text{slide}}$ , measurements of the barrel geometry was needed. Because of the geometry of the wine barrel, barrel sliding causes barrel rotation. The barrel geometry was approximated as a parabola from measured data points. The circumference of the barrel was measured in different locations longitudinally along the length of a typical wine barrel. The data points are provided in Table 4-2 and the parabolic least squares fit is giving in Figure 4-11. Knowing the wine barrel profile allowed calculation of the rotation due to barrel sliding.

**Table 4-2 Barrel Curvature**

Curvature of Barrel			
X from mid	Circumerance	Radius	new axis
(in)	(in)	(in)	(in)
0	85	13.53	0.000
5	83	13.21	0.318
11.5	76	12.10	1.432
16	71	11.30	2.228



**Figure 4-11 Barrel Curvature**

### 4.3.3 Barrel Centroid Location

After establishing the rotation which occurred from rocking and sliding, the location of the centroid of each barrel was calculated. Upon finding the location where a rock corner connects to a barrel, the centroid of each barrel could be found through basic trigonometric relationships.

Figure 4-12 shows a rocking four stack and the method of finding the centroid of the fourth barrel. From knowing the contact points and barrel-rack geometry, the centroid of any barrel can be found at any time step.

Most variables defined for this data reduction are dependent on  $\Delta X_{\text{slide}}$ . Setting up the nonlinear equations in time and using the Matlab function “solve”, the value of  $\Delta X_{\text{slide}}$  was found for each time step and at each barrel height.

A summary of the variables used for data reduction are provided in Table 4-3.



**Figure 4-12 Locating the Centroid**

Name	Dependent On:	Number of Additional New Unknown Variables	Number of Equation	Description
$X_{trans} \# (t)^*$	Known	0	0	Output from the horz. Transducer
$X_{rock} \text{ Box } \# (t)$	$\theta_{rock} \#$ and all below	10	5	Horz. translation of the instrument box due to rocking
$X_{slide} \text{ Box } \# (t)$	$X_{rock} \#, X_{trans} \#$	15	5	Horz. translation of the instrument box due to sliding
$X_{slide} \# (t)$	$X_{slide} \text{ Box } \#, X_{rock} \#$	0	10	Horz. translation of the barrel's centroid due to sliding
$X_{rock} \# (t)$	$\theta_{rock} \#$ and all below	0	5	Horz. translation of the barrel's centroid due to rocking
$\theta_{rock} \# (t)$	$Y_{trans} \#, Y_{error} \#, a \#$	15	5	Rotation of the barrel's centroid due to rocking
$Y_{trans} \# (t)$	Known	NA	NA	Output from the Vert. Transducer
$Y_{error} \# (t)$	$Y_{trans} \#, Y_{trans} \#-1$	0	5	Error in the vert. transducer output due to horz. Translation
$Y_{rock} \text{ Box } \# (t)$	Not Needed	NA	NA	Vert. translation of the instrument box due to rocking
$Y_{slide} \text{ Box } \# (t)$	Not Needed	NA	NA	Vert. translation of the instrument box due to sliding
$Y_{rock} \# (t)$	Not Needed	NA	NA	Vert. translation of the barrel's centroid due to rocking
$Y_{slide} \# (t)$	Not Needed	NA	NA	Vert. translation of the barrel's centroid due to sliding
$r(0)$	Known	NA	NA	Initial distance between rack leg (pivot point) and transducer
$r_i(t)$	$a \# (t)$	5	5	Distance between rack leg (pivot point) and transducer at time t
$a \# (t)$	$X_{slide} \#$	0	5	Horz. distance from rack leg to transducer at time t
$d_{unit}$	Known	NA	NA	$h_b + h_r$
$\theta \# (t)$	$\theta_{rock} \#, \theta_{slide} \#$	10	5	Rotation of individual barrel
$\theta_{Total} \# (t)$	$\theta \#$ and all below	5	5	Rotation of individual barrel + all barrel rotations below
$\theta_{slide} \# (t)$	$f'(x), X_{slide} \#$	0	5	Rotation of the barrel's centroid due to sliding
$f(x)$	Known	NA	NA	Function of the parabola that represents the barrel curve
$X_{Total} \# (t)$	$X_{trans} \#$	5	5	Horz. translation of the of the barrel's centroid
$X \# (t)$	$X_{trans} \#, X_{trans} \#-1$	5	5	Horz. translation of an individual barrel
SUM		70	70	

\*# represents a barrel and is counted from the bottom up  
In Order of Appearance in paper  
Barrel Dimensions are not mentioned

Table 4-3 Variables used in Data Reduction

## Chapter 5

### Analytical Model

---

A computer model of a wine barrel stack was created to gain a broader insight into the seismic behavior of wine barrels. With a wine barrel structural model, measured shake table accelerations from the different experiments were input into the model and nonlinear time history analysis was performed. The data from the shake table tests was then used to verify the validity of the model. The wine barrel stack model was developed in SAP 2000 Advance 9.0.3.

#### **5.1 SAP Advanced Model**

The barrel assembly was modeled as a rigid body (the wine barrel) resting on a flexible rack. For all models, linear elastic modal analysis as well as nonlinear time history analyses was performed. Model calculated periods of vibration and time history response quantities were compared with results from the physical simulation. Elastic models for each barrel stack height were constructed to find the stiffness and modal periods of the various stack heights. The stiffness found from elastic analysis are used to construct appropriate nonlinear friction element input parameters.

The barrel was simulated with six nodes constrained to each other with body constraints enabling translation motion as a rigid body while always experiencing the same magnitude of rotation. Four of the six nodes represented the corners of the barrels and establish the slenderness of the stack. The other two nodes were located at the mid-height of the barrel. An element with a distributed line mass was created between these two nodes (Figure 5-1).

The barrel rack was set up as a series of tubes. The tube used in model creation was a 1.5 inch x 1.5 inch, 0.065 inch thick section consistent with the tested Topco WR2 racks. A36 steel was assigned to the section and moment releases were placed at and barrel connections.

### The SAP 2000 Wine Barrel Model

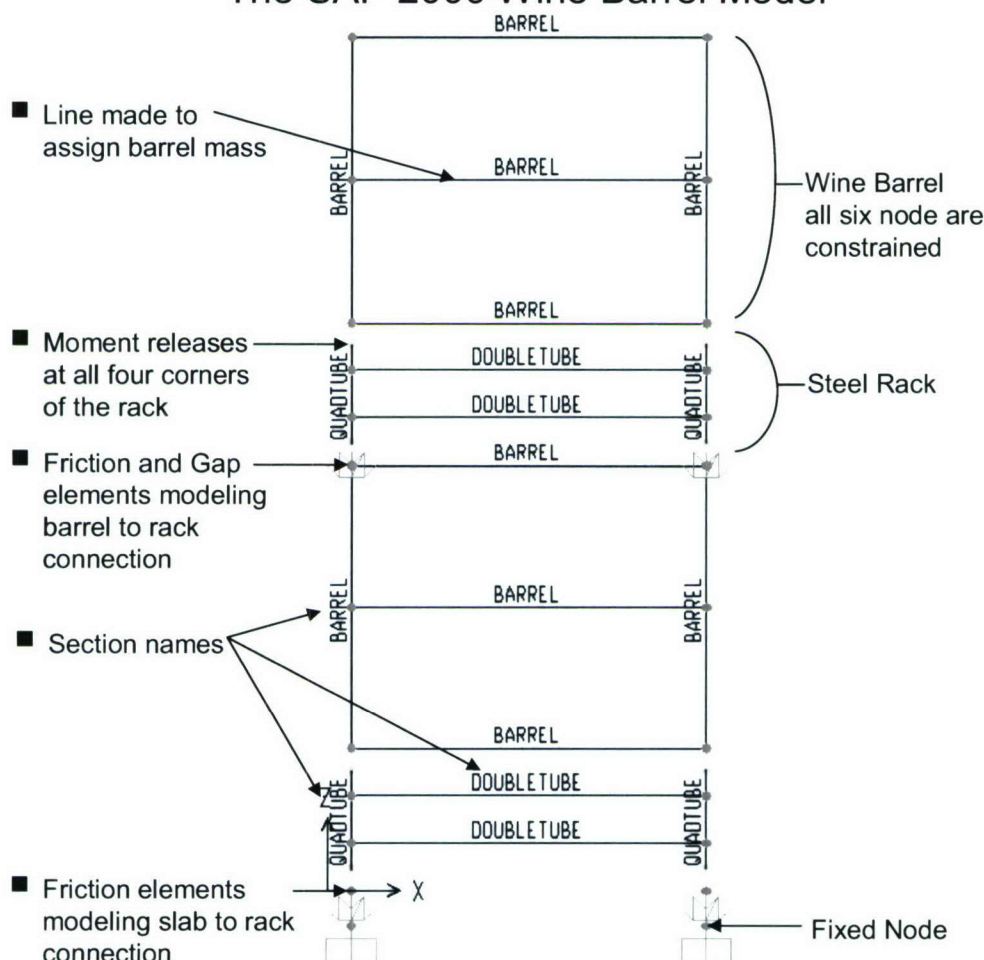


Figure 5-1 Two Barrel Stack Computer Model

Friction pendulum elements were located between the bottom of the rack and either a fixed point (the slab) or the barrel below. These elements were made to capture the energy loss in the seismic activity of the wine barrel and distinguish the model as nonlinear. The friction ratio was established to be 0.47 for rack to slab and 0.30 for rack to barrel contact (Agboatwala, 2005). The inherent in the friction pendulum element is a gap element which enables uplift of the rack corners. A dampening element coincided with the gap element and was set at 2% of critical modal damping. The elastic stiffness in the horizontal direction (U1) was set by the results from the elastic SAP model. The nonlinear stiffness was set two orders of magnitudes smaller than the rack elastic stiffness

to preserve numerical integrity during computation. The axial stiffness in compression was set two orders of magnitude larger for the same reason. The horizontal axis (U3) was fixed because the model is 2D. System damping was set at 2% of critical modal damping.

Nonlinear direct integration time histories were analyzed for each shake table test using the measured table acceleration as input into the computer simulation model.

## Chapter 6

### Analytical and Test Results

The result from extensive experimental testing, consistent with current understanding, is that the portable steel rack method of storing wine barrels is unsafe and not economically sound. In all LA16, most LA18 and some LA19 time histories (representative of the design hazard) the barrel stacks collapse. In this chapter, analysis of the two most common wine barrel failure mechanisms is discussed as well as frequent trends observed throughout testing. Finally, the comparisons between the full scale test and the computer model are made.

#### 6.1 Barrel Mechanisms

The rocking phenomenon can create collapse by either reaching a critical friction angle at a given momentum or by impact force created on the way back (Figure 6-1 and Figure 6-2). The sliding mechanism can also end in a stack collapse as a group of barrels slide off the supporting barrels below. Both of these failures are predominately catastrophic and typically result in complete collapse.

A pure rocking collapse would not normally occur because of the relatively

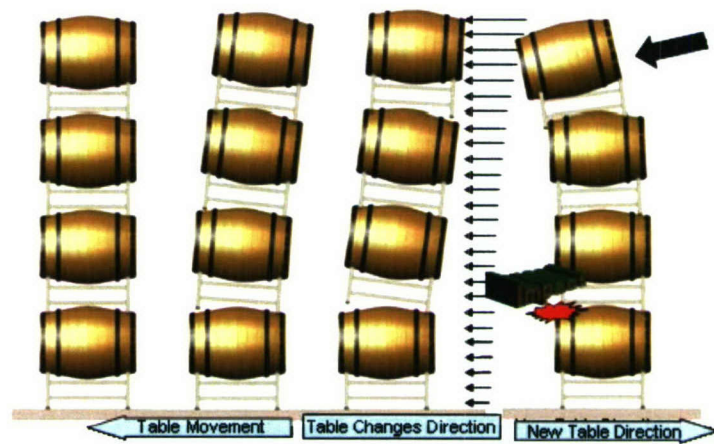


Figure 6-1 Impact Failure

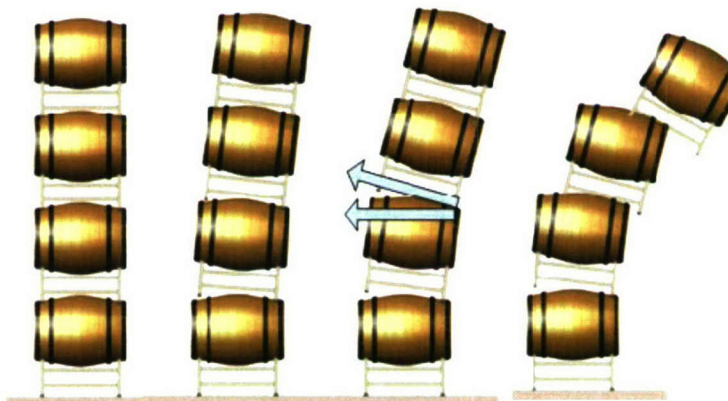


Figure 6-2 Critical Angle Failure

small friction coefficient between the slab and the rack. For any height stack 30 in of pure rocking displacement is needed to create collapse (moving the centroid of the stack past the rotation point). This is roughly three times more than the experimental rocking displacements. The rocking period is defined as the period of a rigid body with a friction coefficient large enough to induce no sliding and pure rocking under seismic excitation first described by Housner, 1965 then again by Markis, 2003.

$$T_{n \text{ Housner}} = \frac{2 * \lambda}{\rho} = \frac{2 * \lambda}{\sqrt{\frac{3 * g}{4 * R}}}$$

In an effort to gain a better understanding into the rocking verses sliding behaviors, a pushover analysis was performed on a rectangular block. There are three stages, which are outlined in figure below which involve in the static capacity of a slender block or simplified wine barrel stack. In the first stage the system deforms elastically and linearly deforms until it begins to rock or slide. The structure will restore to its original shape in the stage. If a force is stepped in a pseudo-static manner it will either rock or slide depending on the block's slenderness.

By finding the sliding capacity of different stack heights and comparing them to the overturning capacity, the point between rocking and sliding can be found. Figure 6-3 shows the rocking/sliding mechanisms dependency on the friction coefficient and slenderness. This is plotted with against each barrel stack's slenderness and a friction coefficient of consistent with steel on wood (0.3). Points above the dividing line will never slide and points below the line slide and then rock. During testing, there were situations where three, four, and five stacks did slide. This would indicate the acceleration step was beyond the capacity of the rocking block.

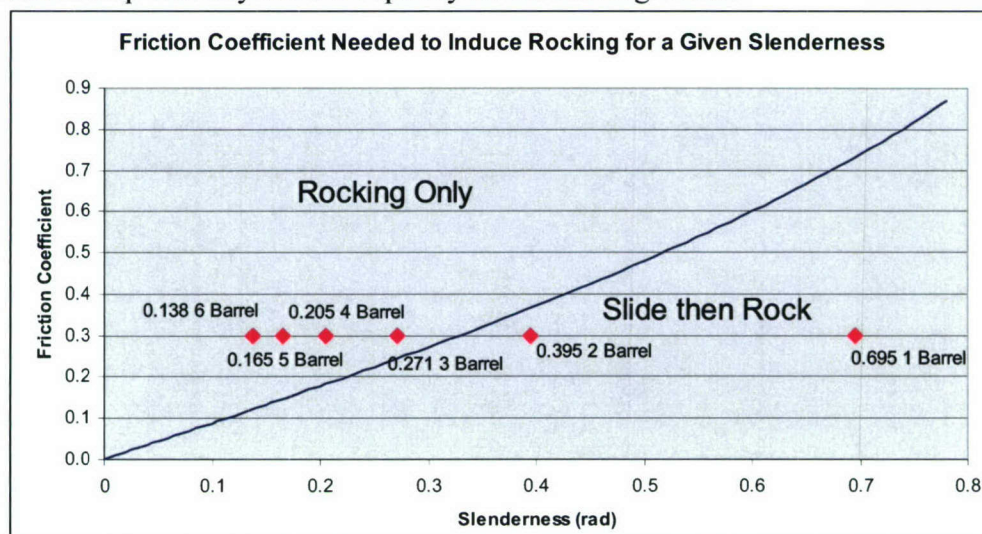
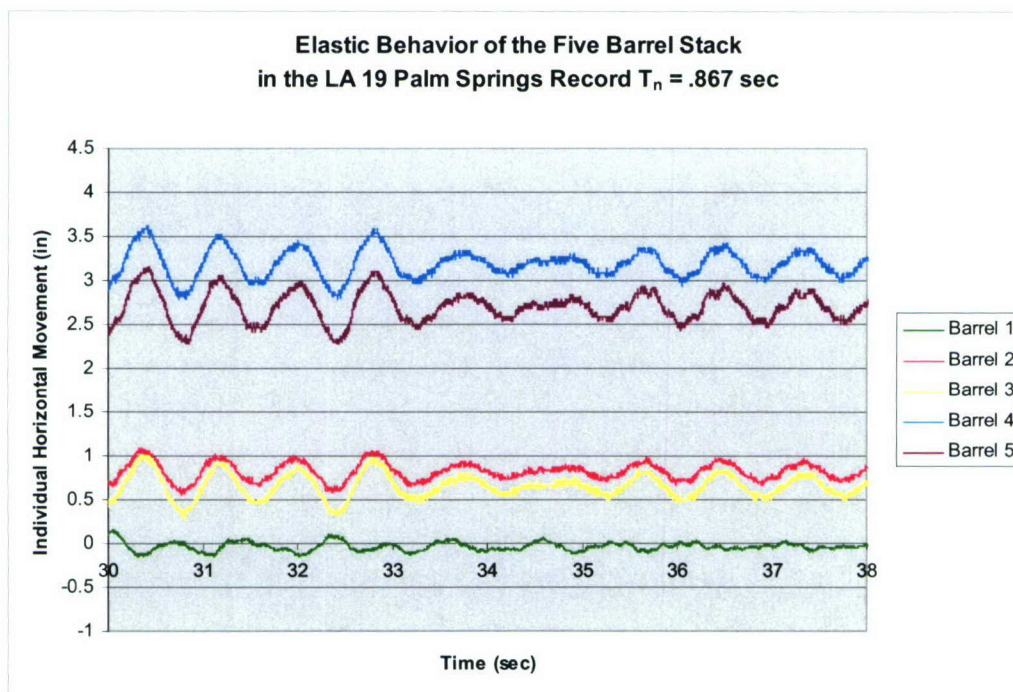


Figure 6-3 Slenderness and Friction Coefficient effect on Barrel Behavior

## 6.2 Experimental Test Results

During LA19 five barrel test, white noise near the end of the experimentation generated a first mode elastic response. The structure moved at roughly the same period with different amplitudes for close to twenty five seconds. The stack also oscillated about the same displacement indicating elastic movement. This elastic response was not observed in the LA 18 (five barrel) or LA16 tests because of high pseudo accelerations for longer periods. These values are two and three times that of the LA 19 response spectra at 0.9 sec.



**Figure 6-4 Elastic Period of the Five Barrel Stack**

While stack collapse did not occur in the later part of the LA 19 record for the one, two, four, and five barrel stacks; a linear relationship in elastic rocking magnitude was found. Each consecutive barrel height moved about 0.3 inches more than the previous height. Earlier in the above graph the fifth barrel oscillated  $\pm 0.8$  inches about the at rest position. The one barrel had  $\pm 0.15$  inches capacity. The elastic period for the five barrel was around 0.867 sec and the four barrel was 0.524 sec.

For all tests, the third barrel never had a noteworthy contribute to a mechanism. The third barrel never initiated uplift or slid more than 0.6 inches (always orders of magnitude less than the surrounding barrels behavior).

### 6.3 SAP Advanced Results

As per chapter five, an elastic model was created to find the periods and stiffness of each barrel stack height. The total stack stiffness was found by taking the inverse of the displacement at the top node where a one kip load was acting. Calculated periods for the various stack heights are provided in Table 6-1.

Calculated periods of vibration for the 4 and 5 barrel stacks did not articulate well when values calculated from the ground motion white noise experimental measurements.

**Table 6-1 Elastic SAP Results**

Stack Height	Period for Mode Shape			Elastic Stiffness (kip/in)
	Mode 1 (s)	Mode 2 (s)	Mode 3 (s)	
1	0.06055			32.68
2	0.09828	0.03630		16.56
3	0.13790	0.04815	0.03221	10.66
4	0.17971	0.06129	0.04184	7.54
5	0.22355	0.07559	0.04642	5.61

Nonlinear time history analyses predict collapse of all 4 and 5 barrel tests as indicative of the non converging analytical solution. Investigation is currently being performed on scaled versions of the table recordings to ascertain the nature of the ill conditioned problem and to establish the limit of the software's ability to capture the solution.

## Chapter 7

### The Seismic Rack

---

Given the proven instability of wine barrels stacking using conventional methodology, Wineries have a need for a safer way to stack wine barrels in California. A better stacking method must be both inexpensive and convenient for it to be practical for the wine industry. The prototype of the seismic rack described herein satisfies both criteria. The main purpose of the seismic rack is to reduce the first mode response associated with rocking and stack collapse thereby decreasing the risk of loss. This chapter explains details of the investigation and testing into the development of the prototype seismic wine barrel rack.

#### **7.1 The Seismic Rack Prototype**

A standard rack (western saddle) was implemented in the prototype as it is a typical basic building block for stacking wine barrels. The bottom legs of the bottom rack (the bent steel plates in contact with the slab) were removed and replaced with ball bearings placed in a Teflon socket.

In prototyping the seismic rack, the design parameters were socket material, ball bearing material and size, number of ball bearings and legs, shape and size of socket. Investigative findings on these parameters are provided in Table 7-1.

The embedment of the ball bearings into the concrete slab was an issue addressed in the process of creating the prototype. Embedment of the ball into the concrete under gravity load would not only be unattractive to the winery's concrete foundation, but would inhibit the balls ability to roll with low friction thus transferring a larger than design lateral force into the wine barrel stack. Using the required concrete bearing area and assuming a rigid ball, the required embedment depth and diameter were found (Table 7-2).

A series of tests were performed using a 1" diameter chrome ball bearing. Loads were applied against a concrete block and diameters were measured. Test results are given in Table 7-3. Theoretical values were 4 to 8 times conservative as compared to the testing.

Table 7-1 Iterating the Seismic Rack

Issue	Notes
<b>Socket Material</b>	
Steel	
Coated with Teflon	Expensive and chance of chipping off
Not coated	Friction coefficient is too high, but very cheap
PTFE Teflon Plate	Superior performance, Only reasonable option
<b>Ball Bearing Material</b>	
Chrome	corrosion issues, but very cheap
Stainless Steel	Not economically feasible
Composite	Not economically feasible
Teflon	Elastic Modulus is too low
Others	To be investigated
<b>Concrete Embedment</b>	
Number of Contact Points	
Size of Ball Bearings	SEE Fig. 6.2
<b>Location of Legs</b>	
Number of Legs	8, 16 (12 To be investigated)
Equal Load Per Leg	Flexible tubing accommodates
<b>Shape of Socket</b>	
Circular	Cheap Common Mill
Same Size as Ball	Bearings were pinched in socket
Larger than Ball	Enable rolling and became preferred
Elliptical or Parabolic	Custom Mill to be made
Bearing on Teflon	Pinched ball after creep (2-4 hours)
Bearing on Steel Plate	Did not experience significant creep
<b>Size of Socket</b>	
Fit in Normal Tubing	Not adequate for bearing on concrete
Fit in Rectangular Tubing	Does not maximize rolling area
Fit in Oversized Tubing	To be investigated
<b>Size of Ball Bearing</b>	
1.25"	Too Large for tube, expensive
1"	Best results
.75"	Could work, but there were rocking issues when tested
7/8"	To be investigated

Table 7-2 Embedment Depths

The Constants	
Weight per Barrel	562.9 lbs.
Weight Per Rack	35 lbs.
Ball Diameter	1.25 in
Concrete Strength	2800 psi

Table 7-3 Parametric Study of Embedment Depths

Parametric Study									
Stack Height	2			4			5		
Balls per Rack	8	12	16	8	12	16	8	12	16
Weight per ball (lbs)	290.2	193.5	145.1	580.4	386.9	290.2	725.5	483.7	362.8
Embedment Diameter (in)	0.363	0.297	0.257	0.514	0.419	0.363	0.574	0.469	0.406
Embedment Depth (in)	0.0259	0.0174	0.0131	0.0507	0.0343	0.0259	0.0628	0.0425	0.0322

A time history analysis, friction slider model was written in MatLab to verify that the slab in the ATL was adequate for testing the seismic rack. The program ran through earthquake time histories (LA16, LA18, and LA19) with different stack weights (associated with the 2, 4, and 5 barrel stacks) and also varied friction coefficients.

The layout of wine barrel stacks in a storage facility would have to accommodate the large displacements of the prototype racking system. A rack that moved more than 18" would not be practical for the cellars. Most wineries have around 6" to 12" of space between wine stacks and around 24" of space between stacks and the wall. Typically, every two or three rows normally there is space for a forklift. The displacements given by the friction slider model are within the stack space of a typical warehouse layout. The results of this program can be used to gain insight into the friction coefficient, but errors will exist because a friction slider was used instead of the complete barrel stack model. It should be noted that larger displacements are associated with less weight or smaller friction coefficient. A summary of the analytical investigation of the seismic rack using LA16 and LA18 ground motion records is provided in Table 7-4.

As is expected, analytical simulation shows that as the friction coefficient of the base rack is increased the sliding displacement is proportionally decreased. In addition, simulation results show that the higher wine barrel stacks tend to slide less for a given constant friction coefficient (Figure 7-1 and Figure 7-2).

Table 7-4 Movement of a Friction Slider

LA 16				
# of Barrels	Stack Weight	Friction Coefficient	Max Displacement	Min Displacement
2	2350	0.100	8.24	-9.21
2	2350	0.050	8.89	-9.65
2	2350	0.030	9.21	-9.86
2	2350	0.010	9.52	-10.06
4	4700	0.100	7.39	-8.54
4	4700	0.050	8.24	-9.21
4	4700	0.030	8.74	-9.56
4	4700	0.010	9.38	-9.96
5	5879	0.100	7.08	-8.28
5	5879	0.050	7.99	-9.02
5	5879	0.030	8.54	-9.42
5	5879	0.010	9.30	-9.91
6	7050	0.100	6.82	-8.05
6	7050	0.050	7.76	-8.85
6	7050	0.030	8.36	-9.29
6	7050	0.010	9.21	-9.86
LA 18				
# of Barrels	Stack Weight	Friction Coefficient	Max Displacement	Min Displacement
2	2350	0.10	4.69	10.69
2	2350	0.05	4.43	11.36
2	2350	0.03	4.97	11.61
2	2350	0.01	6.00	11.67
4	4700	0.10	5.08	9.62
4	4700	0.05	4.69	10.69
4	4700	0.03	4.30	11.23
4	4700	0.01	5.38	11.68
5	5879	0.10	5.13	9.19
5	5879	0.05	4.85	10.40
5	5879	0.03	4.69	11.02
5	5879	0.01	5.16	11.65
6	7050	0.10	5.13	8.80
6	7050	0.05	4.96	10.12
6	7050	0.03	4.62	10.82
6	7050	0.01	4.97	11.61

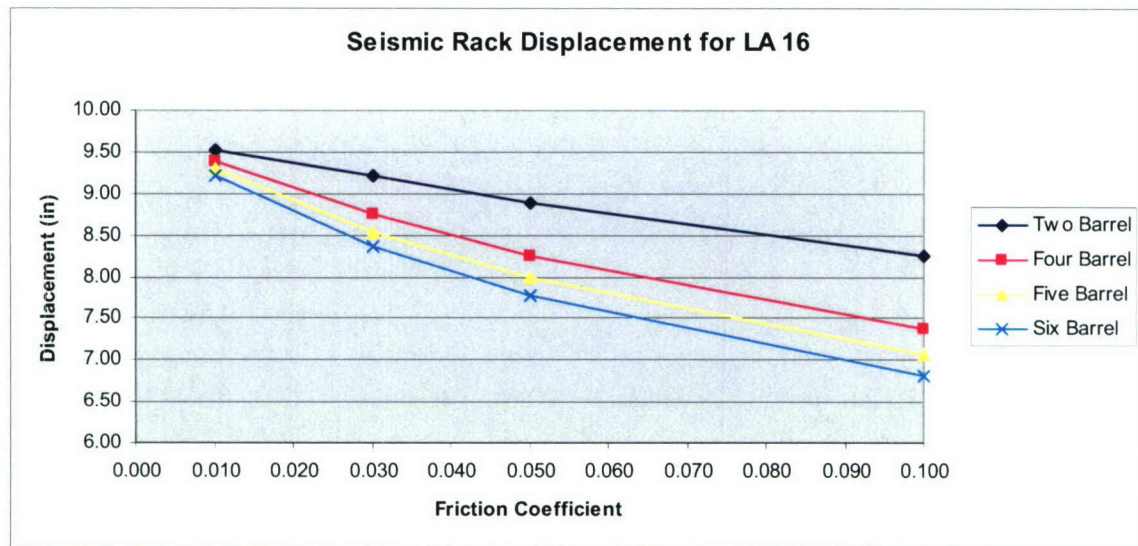


Figure 7-1 LA 16 Displacements vs. Fiction Coefficient

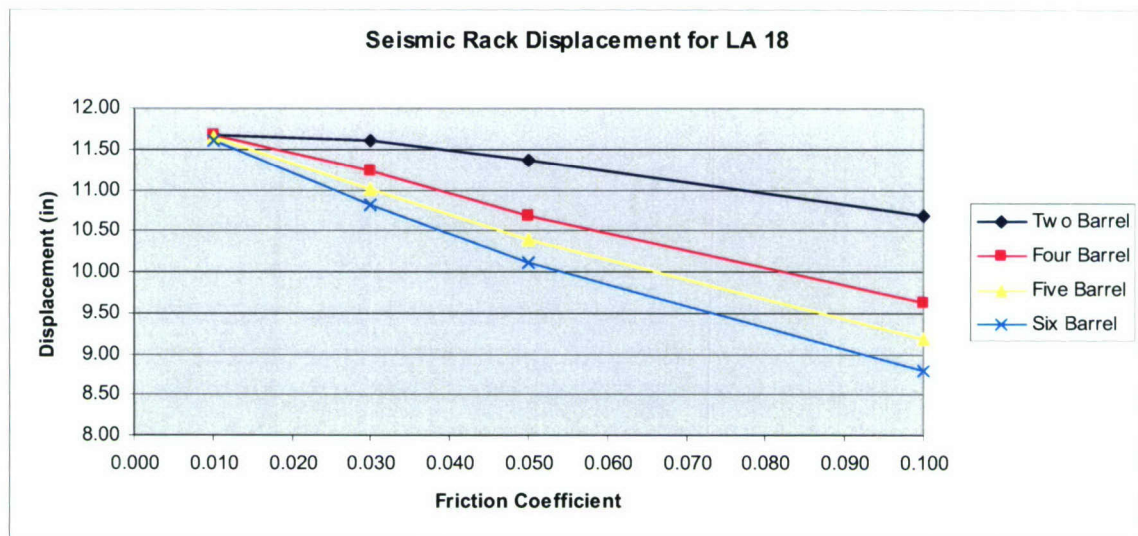


Figure 7-2 LA 18 Displacements vs. Friction Coefficient

## 7.2 Outline for the Prototype Testing

The main source of record keeping and analysis for each test was made by thorough notes. Each test would give insight into making a better rack and normally only one alteration was made to the rack between tests to be conclusive on its contribution. During

prototype investigation, no instrumentation (position transducers or accelerometers) were used but rather observational data was collected through digital video collection and real time surveillance of the tests. Observational data is provided of the prototype testing is given in Table 7-5.

**Table 7-5 Prototype Testing Evolution**

Test #	Stack Height	Ball Size	# of Balls	Date	Records Tested	New To Prototype
1	2	1.25	8	7/23/2006	LA19,LA16,LA18	Initial Rack
2	4	1.25	8	7/24/2006	LA19,LA16	Initial Rack
3	4	1	8	7/25/2006	LA16,LA18,LA19	Smaller Balls
4	4	1	8	7/25/2006	LA16,LA18,LA19	Time for Creep
5	4	1	8	7/27/2006	LA16,LA18,LA19	steel bearing
6	4	0.75	16	7/29/2006	LA16,LA18,LA19	16 Smaller Balls
7	5	1	16	7/31/2006	LA16,LA18,LA19	Bigger Balls
8	4	1	16	8/1/2006	Belford, 16,18,19	New Lubrication
9	4	1	16	8/2/2006	LA16(2x)	Same For Expo

Test #	Comments
1	Although the stack never fell, Slight rocking occurred
2	19 had little difference from normal rack and 16 failed in rocking
3	One of the best tests, No noise and rocking or stack failures
4	Sliding of the forth barrel, no stack failures, serious creep in Teflon
5	Cut deeper into Teflon, Ok test, slab is getting very rough
6	Slight rocking, no bearing issues, good test overall
7	Great test, 2nd to test 3
8	New lube worked better, plus it's more durable, Survived Belford record
9	Good test, Nothing fell!

### 7.3 Analysis of the Testing

The testing of the seismic rack gave insight into the seismic behavior of wine barrels. One key observation was that the overall stack behavior was better when subjected to ground motions with large velocities and small accelerations while the behavior was worse during testing with ground motions containing high acceleration content. The high acceleration type records led to a top barrel walking type mechanism. Without the isolation bearings, the observed stack behavior was exactly the opposite. The high frequency, high acceleration content to the ground motion excites the multi-barrel stack's higher modes leading to small inching type motions of the top barrel. With the fixed base (or steel on concrete bearing), the first mode drove the stack collapse. With the isolation bearings in place, the first mode effectively (through large sliding) is lengthened such that collapse is prevented.

## Chapter 8

### Conclusions

---

Wine barrel stacking methodology typical of California wineries is to vertically stack oak barrels between light gauge portable steel racks. The individual racks have no connection to the ground, each other, or the adjacent barrels. This lack of positive connection, while adequate for gravity loads, is inadequate for lateral loads stemming from seismic ground excitation. The study described in this document assessed the vulnerability of 2 barrel, 4 barrel, and 5 barrels stacks. Results indicate, consistent with past findings post earthquake, that the taller wine barrel stacks are subject to catastrophic collapse. Barrel stacks sitting on a concrete sub-surface are most vulnerable to near fault high velocity type ground motions resulting in a predictable and consistent total stack collapse.

Model simulation did not articulate well with the measured periods of vibration. Furthermore, consistent with the findings from other researchers, analytical modeling of this chaotic nonlinear rocking response proved difficult resulting in a numerically unstable system model. Further research and analysis is needed ascertain the limits of a stable numerical solution of a collapsing wine barrel stack.

Simple analyses based in modal periods and the Housner rocking period did provide insight into the relationship between rocking and sliding as a function of inter-barrel and steel on concrete friction coefficients and overall stack slenderness. Numerical results suggest that lowering the coefficient of friction between the base rack and the barrel stack would result in an increased likelihood of the benign stack sliding mechanism. By changing the barrel stack failure mode to sliding rather than rocking/pounding/barrel ejection, the integrity of the stack can be protected during strong ground shaking.

Given the not unexpected findings from a reduced base rack friction coefficient, a prototype bearing was developed using steel ball bearings mounted inside a spherical pad of Teflon. Preliminary analysis suggested that the coefficient of friction was reduced from 46% to approximately 10%. Analytical results suggest that 5 barrel stacks would slide before rocking for base friction coefficients under 13%. Experimentation of the 4 barrel and 5 barrel stacks bearing on Teflon isolations bearings verified this finding by protecting against stack collapse. All steel on concrete tests of the LA16, and to some extent tests with the LA18 and LA19, ground motion records resulted in stack collapse. With the inclusion of the Teflon bearings in the bottom rack subjected to the LA16,

LA18, and LA19 ground motion records, both the 4 and 5 barrel high stack resulted no stack collapse. It was observed, however, that high frequency content, high acceleration type ground motions did have more of an effect on the top barrel in the stack when the bottom rack was mounted in the Teflon bearings.

The current racking system used in the wine industry has become ingrained into manufacturing processes. The wine industry needs the racks to be vertically unconnected to allow quick movement of the barrels in the stack for vertical rotation of the barrels. Given that vertical anchorage of the high barrel stacks are not an option, reducing the friction the base results in domination of a safer mode of stack movement: sliding rather than rocking. Using isolation bearings on these stacks in laboratory testing has shown this to be a viable solution to the stack collapsing problem.

## Chapter 9

## References

---

- Aslam, M., Godden W. G., Scalise T. D., (1980) "Earthquake Rocking Response of Rigid Bodies", *ASCE Journal of the Structural Division*, V. 106, February, pg. 377-392.
- Beardmore, Roy (2006) "Friction Factors". UK,  
[http://www.roymech.co.uk/Useful\\_Tables/Tribology/co\\_of\\_frict.htm](http://www.roymech.co.uk/Useful_Tables/Tribology/co_of_frict.htm).
- C. B. Chadwell, J.M. Stanley and J. Marrow, (2006) "Nonlinear Analytical Modeling and System Verification Utilizing a Combined Slide-Rock Seismic Response of Wine Barrel Stacks." Earthquake Research Institute, Proceedings of the 13<sup>th</sup> World Conference on Earthquake Engineering, Vancouver, B.C., Canada, August 2006.
- Chopra A.K, (2000) "Dynamics of Structures: Theory and Applications to Earthquake Engineering (2<sup>nd</sup> Ed.)." Prentice Hall: Englewood Cliffs New, NJ, 2000.
- Housner G.W, (1963) "The Behavior of Inverted Pendulum Structures During Earthquakes" *Bulletine of the Seismological Society of America* 1963; 53(2) pg 404– 417.
- Makris, N., C.J. Black, (2001) "Rocking Response of Equipment Anchored to a Base Foundation". PEER 2001/14, Pacific Earthquake Engineering Research Center, University of California, Berkeley.
- Marrow, J.M., (2000) "2000 M5.2 Yountville/Napa Earthquake Implications for the California Wine Industry", Simpson Gumpertz & Heger, Inc., San Fransisco, CA,  
<http://www.vingenuity.net/pdf/2000-yountville-quake.pdf>.
- Marrow, J.M., N. Makris (2000) "Experimental Studies on the Earthquake Performance of Wine Barrel Stacks", Earthquake Engineering Research Institute, Proceedings of the 7th National Conference on Earthquake Engineering, Boston, Mass, July 2002.
- Marrow, J.M., (2002) "The Seismic Vulnerability of the California Wine Industry – An Experimental Assessment", 2002 Structural Engineer's Association of California Annual Convention, Santa Barbara, California.
- PEER, (2003) Pacific Earthquake Engineering Research Center, "Open System for Earthquake Engineering Simulation (Opensees)", <http://opensees.berkeley.edu>.

Psycharis, I. N., Jennings P. C., (1983) "Rocking of Slender Rigid Bodies Allowed to Uplift", Earthquake Engineering and Structural Dynamics, V. 11, Pg. 57-76.

Topco Inc, (2005) "Wine Barrel Rack Specifications", [www.topcoproducts.com](http://www.topcoproducts.com), 974 Griffin St. Grover Beach, Ca 93433.

Wine Institute, (2001) "Commercial Wineries", [http://www.wineinstitute.org/communications/statistics/commercial\\_bondedwinery.htm](http://www.wineinstitute.org/communications/statistics/commercial_bondedwinery.htm).

Wine Institute, (2003) "2002 California Wine Sales", <http://www.wineinstitute.org/communications/statistics/Sales>.

Wine Institute, (2005), "Industry Back ground and Statistics, <http://www.wineinstitute.org>, 425 Market Street, San Francisco, Ca, April 5, 2005.

Winkler, T., Meguro K., Yamazaki F., (1995) "Response of Rigid Body Assemblies to Dynamic Excitation", Earthquake Engineering and Structural Dynamics, V. 24, Pg. 1389-1408.

**The following report has not been included because it contains  
proprietary information.**

### **Three-dimensional scaffolds for mammary epithelial cell growth**

Project Investigators:

Robert Crockett, PI  
Dan Peterson, Co-Investigator  
Rafael Jimenez-Flores, Co-Investigator  
California Polytechnic State University, San Luis Obispo, California

Graduate Students: Megan Barry, Danielle Norris, Meng-Ju (Angela) Hsieh

## **Improving power conversion efficiencies of polymer based solar cells**

Project Investigator:

Robert Echols  
Physics Department  
California Polytechnic State University  
San Luis Obispo, California

## Abstract

Because polymer-based solar cells show great promise as a lightweight low cost method to harvest the Sun's abundant energy, a number of larger companies including General Electric, Shell, Sharp, Sony and Toshiba have begun research programs. However, before polymer-based solar cells become a viable technology, fundamental advancements in power conversion efficiency (light energy into electrical energy) must be realized. During the recent funding period, we have focused on improving power conversion efficiencies of dye-sensitized titanium dioxide ( $\text{TiO}_2$ ) solar cells with PCBM blended poly(3-hexylthiophene) polymer (P3HT) doubling as hole transporter and active layer. We have discovered dramatic device performance improvements resulting from adding a smooth sol-gel  $\text{TiO}_2$  layer prior to depositing the nanoparticle  $\text{TiO}_2$  layer and using an amphiphilic ruthenium dye. Analysis of dark currents and incident photon conversion efficiencies confirm the amphiphilic dyes reduces charge recombination while increasing shunt resistance. Smaller improvements result from blending the polymer with [70]PCBM instead of [60]PCBM. Testing device efficiency dependence on solvent type used to dissolve the PCBM:P3HT polymer blend indicates superior performance from tetrahydrofuran (THF) mixed with toluene. While determining optimal nanoparticle  $\text{TiO}_2$  grain size and layer thickness requires additional research, our best data results from 21 nanometer particle size with 1.2 micron thickness. Lastly, attempts to improve PCBM:P3HT polymer blended devices by doping the conducting polymer (PEDOT) with glycerol proved unsuccessful. We presented our work to an international audience at the Spring 2006 Materials Research Society Conference in San Francisco and have an accompanying manuscript in preparation.

## Introduction

With global energy consumption projected to increase 71% from 2003 to 2030 [1] our finite non-renewable petroleum based resources will continue to dwindle, making the need to develop additional renewable energy sources a necessity for maintaining our current standard of living and improving the standard of living worldwide. In fact, with oil prices surging to near \$80 a barrel recently, signaling the end of cheap oil, many experts believe global reserves are nearly half gone, leading to a future of demand outstripping supply [2]. Solar energy, a readily available underutilized renewable energy resource, currently costs about ten times more than the commercial price of electricity for the average consumer. A major factor contributing to the high cost is the need to use relatively pure crystalline silicon wafers to create long diffusion lengths for conduction band carriers. To circumvent expensive energy intensive purification processing, thin-film solar cell devices using direct band gap materials are currently being developed [3]. A particularly novel class of low cost thin-film photovoltaic devices uses polymers as the light-harvesting medium (for a review see the recent text [4]). While polymer solar cells currently have low power conversion efficiencies (around 5% for the best small scale laboratory cells [5]), new design structures and material improvements will make polymer solar cells a viable alternative in the near future.

Our approach towards improving polymer-based solar cell efficiency draws from two independently successful strategies. The best bulk heterojunction devices blend polymer (usually poly(3-hexylthiophene) or P3HT) with [6,6]-phenyl- $\text{C}_{61}$  butyric acid methyl ester (PCBM) to create a percolating donor-acceptor network between electrodes

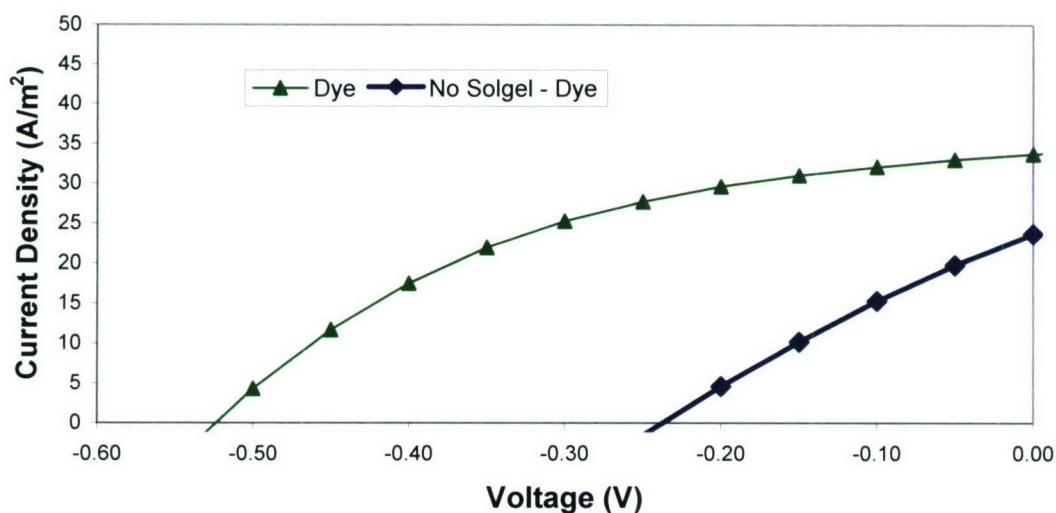
of indium tin oxide (ITO) and aluminum. Solar cells based on the bulk heterojunction concept using the above materials perform near their potential at about 3.0% [6]. However, recent work with a titanium dioxide optical spacer inserted before the aluminum electrode increases the optical field in the cell's light harvesting region, improving power conversion to about 5% [5]. Further improvement will require additional light harvesting capabilities. Our device structure increases light absorption by adding a dye-sensitized titanium dioxide ( $\text{TiO}_2$ ) layer. Alternatively, our design can be viewed as a solid-state Gratzel cell with the polymer blended PCBM acting as hole-transporter and light harvester. Gratzel cell's with liquid hole transporters such as iodine ions convert sunlight to electricity in excess of 10% [7] but suffer from short lifetimes due to electrolyte evaporation. Using solid-state organic molecules and the same amphiphilic ruthenium dye we use, Gratzel recently reports a power conversion efficiency of 4% under standard 1.5 air mass conditions [8]. Gratzel's new device is limited by low hole carrier mobility of the molecules requiring a thinner  $\text{TiO}_2$  nanoparticle layer. In our approach, the polymer blended PCBM hole-transporting layer also acts as an active layer increasing light harvesting potential, allowing for thinner overall device thickness. Below we discuss significant progress made during the past year towards improving power conversion efficiencies of dye-sensitized  $\text{TiO}_2$  solar cells with PCBM blended P3HT polymer doubling as hole transporter and active layer.

### Recent Progress:

#### 1) Adding a smooth sol-gel $\text{TiO}_2$ layer:

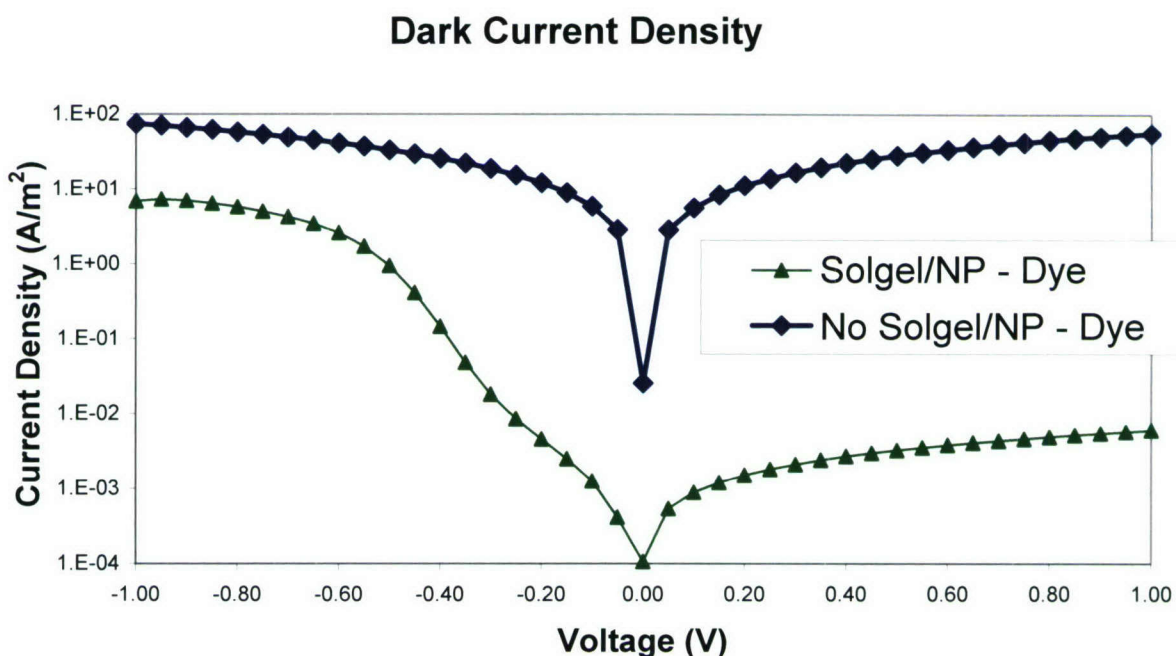
Applying sol-gel  $\text{TiO}_2$  provides a smooth base layer and isolates the ITO electrode from the dye and polymer that penetrates into the nanoparticle  $\text{TiO}_2$  layer

### Current Density Under Illumination



**Figure 1.** Smooth  $\text{TiO}_2$  sol-gel layer dramatically improves open circuit voltage by reducing shorting pathways.

reducing shorts and thus increasing shunt resistance. Substantial increases in open circuit voltage ( $-0.22\text{V}$  increasing to  $-0.52\text{V}$ ) and power conversion efficiency ( $0.18\%$  increasing to  $0.89\%$ ) result for dyed devices as shown in figure 1. Dark current measurements further highlight the effectiveness of the  $\text{TiO}_2$  sol-gel layer towards inhibiting electrons in the ITO from recombining with holes in the dye or polymer blended layers. Figure 2 shows dark current readings for dye-sensitized nanoparticle  $\text{TiO}_2$  solar cell devices under voltage bias ranging from  $-1.0$  to  $1.0$  volts. Under both positive and negative bias the device without the  $\text{TiO}_2$  sol-gel layer allows from over three orders of magnitude more current to one order of magnitude more current at  $1.0$  and  $-1.0$  volt respectively. In addition, little diode rectification occurs for the device without the  $\text{TiO}_2$  sol-gel layer.



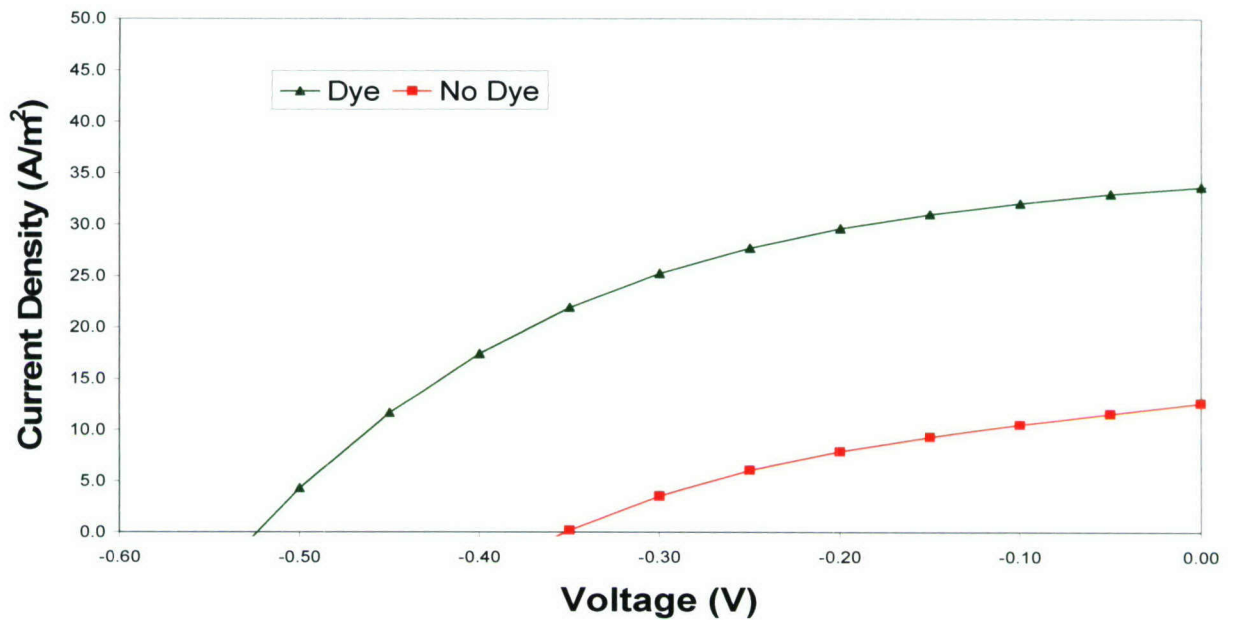
**Figure 2.** Large currents particularly at positive voltage demonstrate the effectiveness of a smooth  $\text{TiO}_2$  sol-gel layer in reducing hole carrier flow from the ITO electrode to polymer blend and dye layers.

## 2) Multiple benefits of amphiphilic dye:

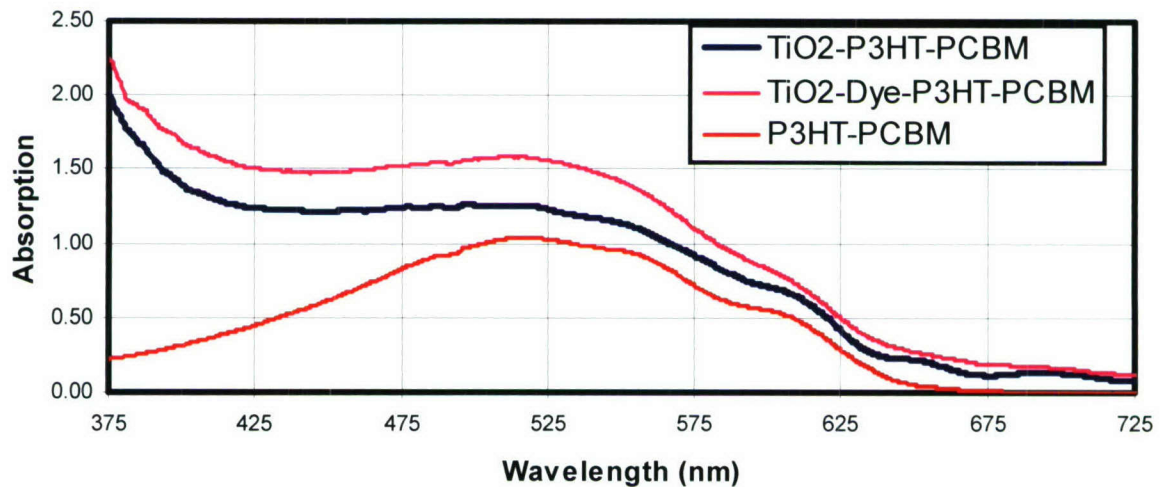
Introducing an amphiphilic ruthenium dyed nanoparticle  $\text{TiO}_2$  layer into a bulk heterojunction P3HT:PCBM solar cell improves absorption and power conversion efficiency. Under  $860\text{ W/m}^2$  white light illumination the dyed device delivers an open circuit voltage,  $V_{oc} = -0.52\text{ V}$ , short circuit current density,  $J_{sc} = 34\text{ A/m}^2$ , and fill factor,  $f.f. = 43\%$ , for an overall power conversion efficiency,  $\eta_{Dye} = 0.89\%$ . An identical device without dye yields substantially lower numbers,  $V_{oc} = -0.33$ ,  $J_{sc} = 13\text{ A/m}^2$ , and  $\eta_{NoDye} = 0.19\%$ , with  $f.f. = 37\%$ . (See figure 3).

Increased photon absorption from the dye proves insufficient to explain the greater than fourfold increase in output power. As shown in figure 4, the dye provides an

## Current Density Under Illumination



**Figure 3.** Dyed cell shows dramatic improvement in current density and open circuit voltage resulting in over 4 times more output power.

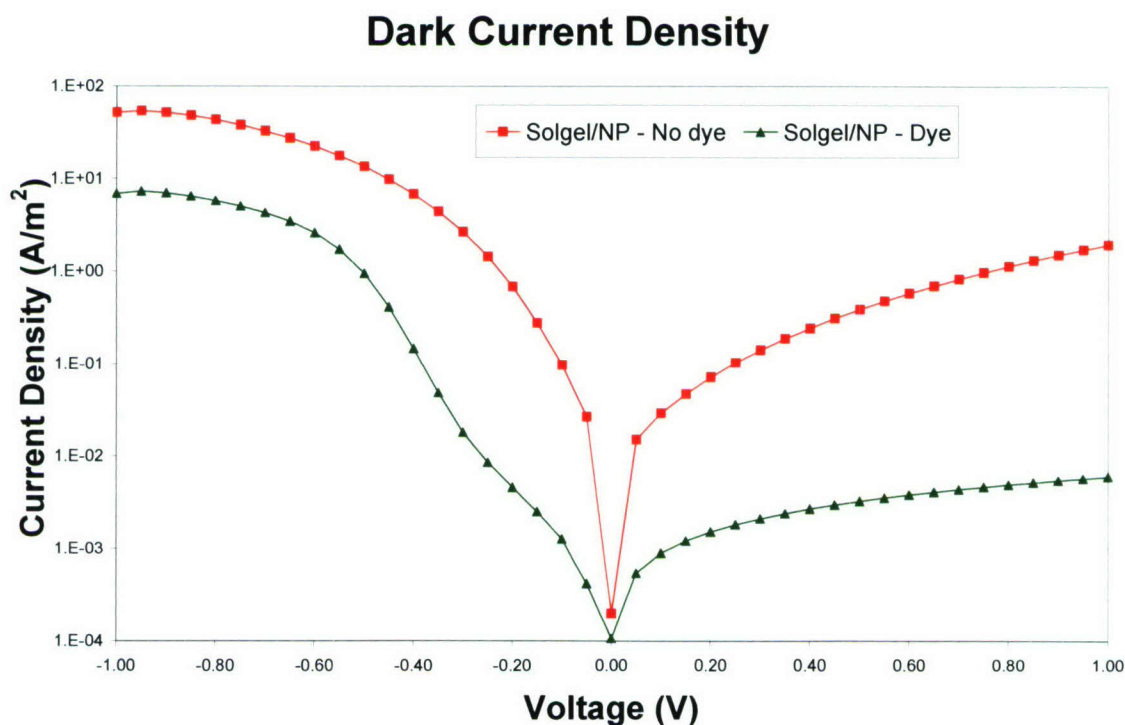


**Figure 4.** Absorption spectra for various cell layers, P3HT:PCBM (red), TiO<sub>2</sub>/P3HT:PCBM (blue), and dyed TiO<sub>2</sub>/P3HT:PCBM (pink).

average 22% absorption increase over the wavelength interval, 375 nm to 725 nm. Assuming an exponential decay in light intensity with depth, the maximum 26%

absorption increase at 520 nm results in a 13% rise in absorbed photons for a one optical depth device. If the solar cell optical depth reduces to one half, 19% more photons are captured. Clearly the increased photon absorption cannot account for all improved power conversion efficiency,  $\eta_{\text{Dye}}/\eta_{\text{NoDye}}=0.89\%/0.19\%=4.7$ . We note that previous work using dye-sensitized  $\text{TiO}_2$  /polymer-only (without PCBM) devices yield overall energy efficiencies of 0.15% [9].

In addition to improving absorption the amphiphilic dye acts as a charge-blocking layer. Figure 5 demonstrates an order of magnitude reduction in dark currents for a dyed device compared to an identical solar cell without dye. Hydrophobic chains on the dye isolate the  $\text{TiO}_2$  layer from the polymer reducing charge transfer. When illuminated, the dye reduces the recombination of  $\text{TiO}_2$  electrons with holes in the dye or polymer, increasing output device current and open circuit voltage. As shown by previous research [10], favorable interactions between the dye hydrophobic tails and polymer increase

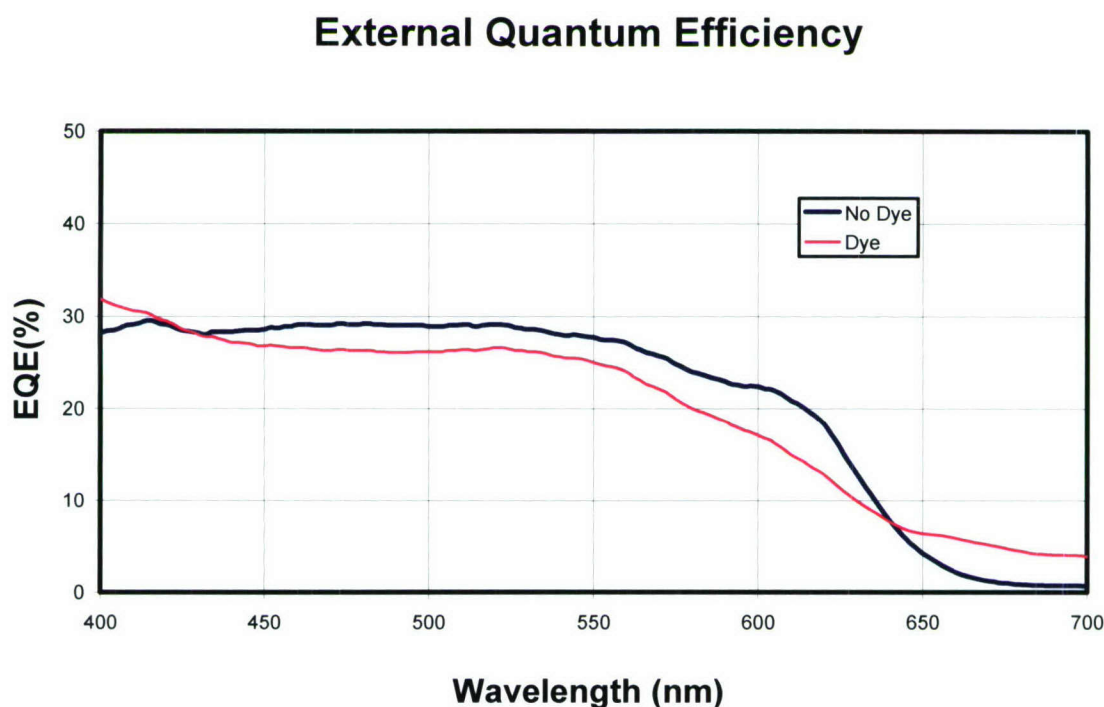


**Figure 5.** Current/Voltage curves without light showing dye lowering currents and increasing diode rectification.

polymer penetration into the  $\text{TiO}_2$  nanoparticle layer. Increased surface area contact between the polymer and  $\text{TiO}_2$  layer improves dye electron regeneration.

To further confirm the amphiphilic ruthenium dye's beneficial role in device performance we conducted external quantum efficiency (EQE) experiments (also known as incident photon conversion efficiency). Initially we expected to perform EQE experiments for short testing times without a protective inert glovebox atmosphere. However, we soon realized device performance degrades rapidly in air. While we are currently working on a fiber optic feed-through system, preliminary results from EQE experiments conducted by very carefully shining light into the glovebox onto devices

yields results shown in figure 6. Interestingly, the EQE for a dye-sensitized TiO<sub>2</sub> nanoparticle device performs worse at almost all wavelengths compared to one without dye. Consequently, one would naively suspect a dyed device to have slightly worse J/V light curves unlike what is shown in figure 3. The discrepancy between the EQE curves and light curves shows that the dye must play a large role in reducing recombination losses at high light intensity by limiting electrons in the TiO<sub>2</sub> or ITO from combining with holes in the polymer or dye as discussed earlier. Figure 6 does, however, show the dyed solar cell performing better at long wavelengths consistent with the absorption spectra in figure 4. An amphiphilic ruthenium dye with even stronger absorption properties in red light has been developed and will be available for general use soon [11].



**Figure 6.** External quantum efficiency (EQE) showing the effects of dye on solar cell performance.

We will be submitting our results for publication after reproducible EQE experiments are fully implemented in the protective glovebox atmosphere.

### 3) [70]PCBM vs. [60]PCBM:

The asymmetrical [70]PCBM fullerene derivative provides more absorption in the visible spectrum compared with the symmetrical [60]PCBM. Previous research confirms the improved [70]PCBM absorption translates into increased power efficiencies[12]. We demonstrate a similar improvement with our dye-sensitized TiO<sub>2</sub> solar cell design. On average the [70]PCBM performs 18% better than [60]PCBM.

#### 4) Solvent type and device performance:

Better performing bulk heterojunction P3HT:PCBM polymer blend solar cell devices result from using solvents in the chlorobenzene family [5]. However, chlorobenzene is the most toxic common polymer solvent. Experimenting with less toxic solvents, toluene mixed in a 1:2 ratio by volume with tetrahydrofuran (THF), xylene, and anisole, we find THF to be superior. The best performing devices for THF, xylene, and anisole yielded power conversion efficiencies of 0.89%, 0.10%, and 0.086% respectively.

#### 5) TiO<sub>2</sub> nanoparticle grain size and layer thickness:

Before experimenting we thought a smaller nanoparticle grain size would lead to smoother films with improved performance. Using 9 nm, 13 nm, and 21 nm average grain size nanoparticle, we found power output from devices with 21 nm TiO<sub>2</sub> nanoparticles to more than an order of magnitude better. Presumably, the larger grain size allows better penetration of the polymer blend, improving contact for electron transfer.

Spinning nanoparticle from a colloidal suspension at one thousand (1k) revolutions per minute (r.p.m.) results in a nanoparticle layer thickness of 1.2 microns measured with an atomic force microscope. At faster spin speeds ranging up to 8k r.p.m. the layer thickness decreases along with solar cell performance. Applying two layers spun at 1k r.p.m. results in a thicker film but reduced device power output. While our best device arises from a 1.2 micron nanoparticle layer, we need to further probe slightly thicker layer performance.

#### 6) Glycerol doped PEDOT:

We attempted to improve performance of a bulk heterojunction P3HT:PCBM polymer blend device with lithium fluoride cathode by doping the PEDOT anode using glycerol to improve layer conductivity. A PEDOT anode without doping remained the best performing device during this testing with  $J_{sc} = 43 \text{ A/m}^2$ ,  $V_{oc} = 0.57\text{V}$  and f.f.=36% for an overall power conversion,  $\eta = 0.98\%$ . We suspect a decrease in layer uniformity for the doped PEDOT resulted in poorer-performing devices.

### Conclusion

Our dye-sensitized TiO<sub>2</sub> solar cells with PCBM blended P3HT polymer doubling as hole transporter and active layer show promise by displaying dramatically improved performance compared to previous dye-sensitized TiO<sub>2</sub>/polymer-only devices. By utilizing a smooth sol-gel layer and an amphiphilic ruthenium dye we demonstrate power conversion efficiencies near 1%. To further improve device performance we will optimize the current design with respect to solvent type, polymer and TiO<sub>2</sub> layer thickness, PCBM concentration and annealing conditions. We expect more dramatic improvements to result from the following modifications: reversing the design architecture as discussed in [5], replacing the P3HT polymer with TIPS-pentacene (an organic soluble molecule with good hole carrier mobility), and using a ruthenium dye with greater absorption over the solar spectrum. Additional experiments will involve replacing a random TiO<sub>2</sub> nanoparticle layer with an ordered mesoporous TiO<sub>2</sub> morphology [11]. We are excited about the distinct possibility that these new design ideas will bring polymer-based solar cells closer to becoming a viable energy technology for the future.

## References Cited:

- [1] The US Department of Energy report, "International Energy Outlook 2006," is available at <http://www.eia.doe.gov/oiaf/ieo>.
- [2] D. Goodstein, *Out of Gas: The End of the Age of Oil*, Norton, (2004).
- [3] A. Shah, P. Torres, R. Tscharnner, N. Wyrsh, and H. Keppner, "Photovoltaic Technology: The Case for Thin-Film Solar Cells," *Science* **285**, 692 (1999).
- [4] S. Sun and N. S. Sariciftci, *Organic Photovoltaics: Mechanisms, Materials, and Devices*, CRC Press, 2005.
- [5] J. Y. Kim, et al., "New Architecture for High-Efficiency Polymer Photovoltaic Cells Using Solution-Based Titanium Oxide as an Optical Spacer," *Adv. Mater.* **18**, 572 (2006).
- [6] R. Echols and C. France (Cal Poly Physics Major), "Theoretical Investigations of Polymer Based Solar Cells," *Mat. Res. Soc. Symp. Proc.*, **822**, S7.7.1-S7.7.6 (2004).
- [7] M. Gratzel, *Nature*, **414**, 338 (2001).
- [8] L. Schmidt-Mende, S. M. Zakeeruddin, and M. Gratzel, "Efficiency improvement in solid-state-dye-sensitized photovoltaics with an amphiphilic Ruthenium-dye," *Appl. Phys. Lett.* **86**, 013504 (2005).
- [9] D. Gebeyehu, C. J. Brabec, F. Padinger, T. Fromherz, S. Spiekermann, N. Vlachopoulos, F Kienberger, H. Schindler, and N. S. Sariciftci, "Solid state dye-sensitized TiO<sub>2</sub> solar cells with poly(3-octylthiophene) as hole transport layer," *Syn. Met.* **121**, 1549 (2001)
- [10] G. P. Bartholomew and A. J. Heeger, "Infiltration of Regioregular Poly[2,2'-(3-hexylthiophene)] into Random Nanocrystalline TiO<sub>2</sub> Networks," *Adv. Func. Mat.* **15**, 667 (2005).
- [11] M. Zukalova, et al., "Organized Mesoporous TiO<sub>2</sub> Films Exhibiting Greatly Enhanced Performance in Dye-Sensitized Solar Cells," *Nano Lett.* **5**, 1789 (2005).
- [12] R. A. J. Janssen, J. C. Hummelen, and N. S. Sariciftci, "Polymer-Fullerene Bulk Heterojunction Solar Cells," *Mat. Res. Soc. Bull.* **30**, 33 (2005).

## Publications and Presentations Related to Project

### Abstracts

*Abra, Lewis; Echols, Robert S.; Hiatt, Michael; 'Dye-Sensitized TiO<sub>2</sub> Solar Cells with P3HT Polymer Blended PCBM as Hole Transporter and Light Harvester'; Poster Presentation to Materials Research Society; April 18, 2006.*

We investigate the behavior of dye-sensitized TiO<sub>2</sub> solar cells with [60]PCBM blended poly(3-hexylthiophene) polymer (P3HT) doubling as hole transporter and active layer. The design consists of PCBM blended in a 4:1 ratio with the polymer P3HT, between a gold cathode and either a nanoparticle or sol-gel TiO<sub>2</sub> anode mounted on an ITO patterned glass substrate. To further improve absorption we dyed the TiO<sub>2</sub> layer with a ruthenium dye. Under 860 W/m<sup>2</sup> white light illumination, the smooth sol-gel TiO<sub>2</sub> device delivers a -0.65V open circuit voltage, 7.2 A/m<sup>2</sup> short circuit current and 21% fill factor for a power conversion efficiency of 0.11%. The nanoparticle TiO<sub>2</sub> device performs similarly with a smaller open circuit voltage, -0.50V, and short circuit current, 3.8A/m<sup>2</sup>, but higher fill factor, 56%, yielding an efficiency of 0.12%. Compared to a previous study using a similar polymer, poly(3-octylthiophene) without PCBM, we have improved the sol-gel device by a factor of 10 by including PCBM. However, we see no device performance change by including PCBM for the nanoparticle material. Since these devices have not been optimized with respect to thickness, PCBM:P3HT ratios, annealing times and solvent type, we expect additional future improvements in power conversion efficiency. We also suspect the nanoparticle device is performing below its potential because of partial shunting across the TiO<sub>2</sub> layer. We are currently working on techniques to remedy this deficiency.

**Development of compaction procedures for optimum placement of  
wastes in municipal solid waste landfills using GPS**

Project Investigator:

James Hanson, Ph.D., PE  
Civil and Environmental Engineering  
California Polytechnic State University  
San Luis Obispo, California

## **Abstract**

This study was conducted to develop procedures for optimization of waste compaction at Municipal Solid Waste (MSW) Landfills. Optimization using GPS technology can improve waste containment efficiency, resulting in conservation of land and environmental resources. A fundamental understanding of waste compaction processes or widely acceptable guidelines for industry use are currently not available. A field study was conducted to identify the baseline parameters and conditions that control compaction of wastes. Field compaction analyses were conducted in a test plot that was part of an active landfill cell at a MSW facility. A GPS system was used to make the required geometric calculations for determination of in field densities and placement efficiency of wastes in the test program. Integration of the GPS system allows for determination of geometric constraints in a manner that has not been previously possible. The study was conducted in collaboration with Riverview Land Preserve (a municipal government operated, MSW landfill in Riverview, Michigan). The landfill operator contributed significant resources to the study including allocation of the test plot, use of compaction equipment, and field personnel to assist with compaction and data collection. In particular, a GPS setup was available at the landfill and was used in the study. Results of the field experiments indicate that compaction of wastes is highly variable during normal operation at a landfill. Measured compacted dry unit weights ranged from 0.9 to 8.0+ kN/m<sup>3</sup> (average 4.8 kN/m<sup>3</sup>). No relationship between compacted dry unit weight and temperature was observed during evaluation of normal operations at the landfill. A strong relationship existed between moisture content and compacted dry unit weight. Optimal moisture content of approximately 65% resulted in a maximum dry unit weight of approximately 8.5 kN/m<sup>3</sup>. Much lower dry unit weights (4 kN/m<sup>3</sup>) resulted for both dry of optimum and wet of optimum conditions. Results from laboratory tests indicated a similar fundamental trend and lower compressibility of wastes was observed for optimum compaction conditions. Significant cost savings and environmental conservation can be achieved by manipulating moisture content during waste compaction.

## **Significance**

Landfilling is the most common means for disposal of wastes in the United States. Municipal solid waste generation in the U.S. has steadily increased from approximately 88 million tons in 1960 to approximately 236 million tons in 2003 (EPA 2005). Despite improvements in source reduction and recycling of wastes, the quantities of wastes disposed in landfills remain high in the U.S. at approximately 131 million tons in 2003 (EPA 2005). While the amount of wastes increased in the last several decades, the number of landfills has decreased due to complexities involved in siting new landfills in greenfield sites and also due to the shift in the industry towards larger landfills in comparison to smaller facilities. Optimization of waste placement is essential in larger landfills to dispose the maximum amount of wastes to prolong the active life of a facility. Sustainable operation of landfills includes optimum management of existing sites as well as prevention of development of new landfills in greenfield sites.

Compaction consists of mechanical densification of geo-materials. It is a fundamental property of geo-materials that affects geomechanical stability, compressibility, and fluid transport (Holtz and Kovacs 1981). Wastes are placed in landfills using compaction procedures similar to those used for other geo-materials such as soils. Compaction controls the amount of wastes that can be disposed at a landfill. Waste containment efficiency at a facility is dependent on the compaction of wastes (Tchobanoglous et al. 1993). Optimum waste containment is

achieved by maximized in-place density of wastes. Maximum amount of wastes are placed over minimum land area footprints of landfills under fully optimized conditions, which result in increased lifetime and conservation of land and environmental resources. Maximum densities also result in improved engineering properties of wastes including increased mechanical stability (resulting in improved slope stability), low compressibility and waste settlements, and low fluid transport. Compaction and resulting waste densities also affect biochemical processes and decomposition of wastes. Generation and mobility of the three significant landfilling byproducts resulting from decomposition and degradation of waste (leachate, gas, and temperature), are all influenced by compaction. Efficiency of alternative energy production at landfills through the use of landfill gas are influenced by compaction conditions. In addition, the densities of as-placed wastes affect air emissions and odor control; affect efficiency of leachate recirculation and other bioreactor landfill operations; and overall affect long-term airspace utilization. Improving the degree and consistency of waste compaction will enhance performance predictions for landfills and allow for better site management for selected design objectives.

Reported unit weights of compacted wastes cover a wide range from 3 to 13 kN/m<sup>3</sup> (Sharma and Lewis 1994, Gabr and Valero 1995, Bleiker et al. 1995, Qian et al. 2002). This large range demonstrates that variables related to compaction method can have significant influence on the volume that each ton of waste occupies in a landfill. However, systematic scientific analyses are not available for parameters such as type of compactor, applied compactive effort, and moisture; and conditions such as climatic region and waste placement sequence, all of which affect waste compaction. Landfill owners and operators typically maintain records of placement of wastes at their sites. In general, these records include some estimate of placement efficiency of wastes. However, variations in as-placed densities or factors that affect these densities are typically not available. In addition, the accuracy and representativeness of the estimates are oftentimes questionable due to the approximate methods used in the estimations. Preliminary site records from the industrial partner in the proposed study indicated that daily compacted waste unit weights vary considerably, from approximately 5.8 to 9.3 kN/m<sup>3</sup> for conditions when one and two compactors are used for waste placement operations, respectively (Bobeck 2004). The site manager has also observed improvement of waste compaction during rainy weather, although no data was available prior to this investigation for quantifying the degree of improvement or the amount of water added to the wastes.

Even though the techniques for soil compaction were standardized in the 1930s, a similar concept has not yet been applied to waste compaction despite the large quantities of wastes disposed in landfills. Two main complications exist that have prevented an effective, systematic investigation of waste compaction from being completed:

- 1) A field-scale compaction test program is necessary to provide representative measurements for wastes, which are characterized by large particle sizes and high levels of heterogeneity. Field studies are typically more costly in comparison to laboratory studies. In addition, full cooperation from a site is required that generally is not readily obtained as the requirements of a scientific investigation typically do not agree with daily operations.
- 2) Precise monitoring of volumes of wastes placed has historically been conducted by topographic survey crews and is generally labor-intensive, relatively expensive, and not necessarily fully representative. Inaccuracies in determination of densities result from the low resolution measurement of large volumes in the field.

In general, systematic field studies have not been conducted to investigate the factors that influence compaction of wastes at MSW landfills. This study was conducted to provide a systematic field study to define the baseline parameters and conditions that control compaction of MSW.

A concept known as “intelligent compaction” has recently developed in the construction industry for earthwork (Briaud 2005). The concept uses real-time monitoring of compaction efficiency and establishes the compaction parameters (number of passes with a compactor and moisture content) to obtain the desired engineering properties in the soil instead of simply targeting maximum density. The underlying principles of the intelligent compaction concept were applied to wastes in this investigation. Compaction activities at the site were monitored throughout the entire duration of the study. An analysis was conducted initially to fully integrate the GPS system to field compaction operations at the site. Then, the effects of various parameters and conditions on waste compaction were investigated using a large field test plot. Finally, laboratory tests were conducted to investigate the engineering performance (notably, compressibility) of compacted wastes as a function of moisture content.

## **Experimental Test Program and Results**

### *Overview*

The test program consisted of field and laboratory testing to evaluate waste compaction as a function of operational and environmental conditions. Field tests were conducted to evaluate waste compaction effectiveness during normal operation at Riverview Landfill. Subsequently, large-scale controlled test plots were used to evaluate the influence of moisture content on waste compaction. Laboratory tests were used to determine fundamental baseline trends on manufactured wastes and to determine the influence of moisture content on the compressibility of wastes.

Initially, an analysis was conducted to fully integrate the GPS system to field compaction operations at the site. Using computerized analysis, specification of real-time modifications to daily compaction practices (including adjustments of compaction effort and moisture content) can be possible. Daily records of compaction parameters and conditions were recorded and transferred to the research team at Cal Poly. Compacted unit weights as determined using the GPS system were monitored throughout the entire duration of the study. Additional data such as climatic conditions (temperature and precipitation), waste stream, and waste placement specifics (locations near the edge of a cell, near the center of a cell, etc.) were also obtained. Correlations were established between the properties of compacted wastes and compaction conditions. These data supplement the information obtained from more systematic studies of compaction at the site using field tests as described below.

A significant portion of the test program consisted of field testing. The industrial partner in the project fully cooperated in conducting this extensive test program. Riverview Land Preserve provided site access to an active landfill cell, field compaction equipment and other heavy construction equipment (including backhoes, water trucks, and waste haulers), and onsite personnel for conducting the experiments. In addition, the site manager agreed to use compaction conditions (such as low compaction efforts or excessive addition of moisture in the

wastes) during portions of the test program that may be counterproductive for obtaining optimal compaction at the site to benefit the scientific integrity of the proposed study. The field tests were conducted in a test plot located within a cell active at the time of testing. The approximate dimensions of the test plot were 6 m x 25 m. Tests were conducted by modifying compaction techniques using the normal incoming waste stream and monitoring the resulting properties of the compacted wastes.

Riverview Land Preserve has provided site access to active cells at their landfill. Four visits to the partner landfill have been made to date for the investigation. Three of the visits were related to planning activities and the other visits were more intensive experimental investigations. Baseline conditions at the site were established. Data were collected on the as-placed temperature of the waste, the incoming moisture content of the waste, and the compacted unit weight of the waste. In addition, electronic files and hardcopy documents related to the investigation were collected. The landfill is a subtitle D municipal solid waste landfill located in Riverview, Michigan. The total permitted airspace (design volume) is 30,470,000 m<sup>3</sup>. The average daily waste intake is approximately 3,000 t/year. The fraction of design footprint covered with waste is currently 91%. The design waste placement area is 104.5 ha.

The Global Positioning System at the site was offline for extended periods due to a conflict of radio frequency transmission with the Canadian government. The site is located within approximately 4 km of Canada. Approval for a new broadcast frequency was required by both the United States FCC and by the Canadian government. Approval was requested in September 2005. Approval was granted for a new frequency in May 2006. During May and June 2006, some technical difficulties were encountered bringing the system back on line. Baseline topographic elevations at the site needed to be re-established for any region of the landfill that had a net change in elevation due to waste placement or settlement since the GPS system was offline.

#### *Moisture Content*

The gravimetric moisture content,  $w$ , at which waste arrives to the landfill is required in this study. The procedure used to determine the moisture content of the incoming waste was based on ASTM D2216-05: Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass. The gravimetric moisture content of 28 samples was determined. Most samples represented a miscellaneous mix of waste. Some samples were collected representing specific materials such as construction and demolition material and wood chips that are used for cover at the landfill. The waste used for the samples was delivered to the shop in the bucket of a front end loader operated by an operator at the landfill. Sample sizes ranged from 1 kg to 3.2 kg with the majority of samples being between 1 kg and 1.5 kg. Wastes were dried in ovens at 105°C. The moisture content values determined from waste samples varied significantly. This can be attributed in part to the heterogeneity of the materials present in the samples. Overall, moisture content was determined to range between 6.4% and 62.5%. Average values of moisture content based on waste type is presented in Table 1. The moisture determinations indicated that the waste contained higher moisture contents in summer (June) than in winter (December), which was consistent with general observations of field operations. Frequency histograms of the sampling periods are presented in Figure 1.

Table 1. Summary of Moisture Content Determinations

Waste Type	Moisture Content
Miscellaneous	33.0%
Construction and Demolition	11.6%
Household	27.5%
Wood Chips	46.1%

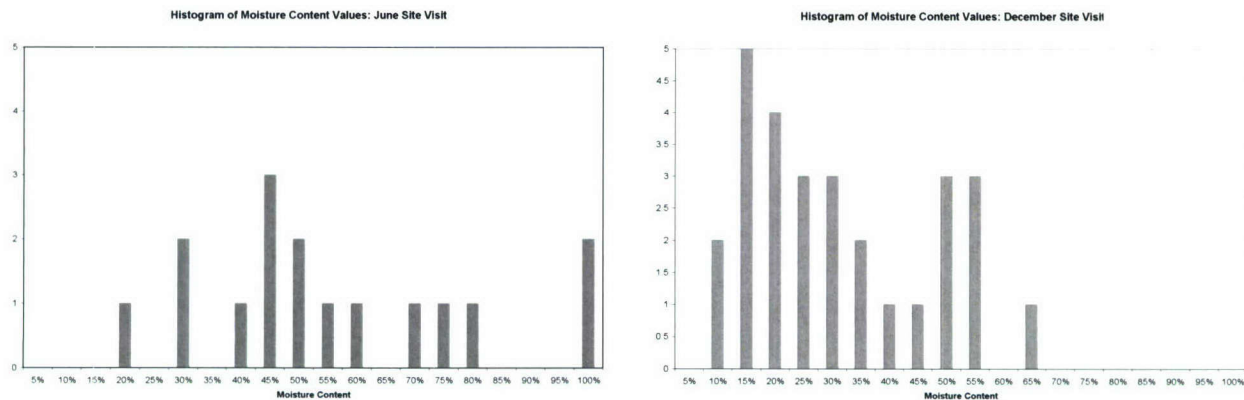


Figure 1. Frequency Histograms of Moisture Content Determinations of Incoming Wastes

### Surface Temperature

A summary of the waste temperature data collected is presented in Table 2 and frequency histograms are presented in Figure 2. For winter conditions, the average recorded temperatures are generally consistently lower than the corresponding air temperature. For winter conditions, the average recorded temperatures are closest to the average daily low temperatures for the vicinity as determined using local climate data. For summer conditions, the average recorded temperatures approach average daily air temperatures. This distinction has implications for predicting waste temperatures during compaction from air temperature data.

Table 2. Summary of Surface Temperatures and Air Temperature Data

	Waste Surface Temperatures				Scale House	NOAA's Weather Service		
Date	No. of Readings	Average	Min	Max	Air	High	Low	Average
12/14/2006	60	-6.5 °C	-7 °C	-4 °C	-2 °C	-1 °C	-8 °C	-4 °C
12/15/2006	82	-5.9 °C	-8 °C	-4 °C	1 °C	1 °C	-2 °C	0 °C
12/16/2006	61	-5.4 °C	-8 °C	-1 °C	-2 °C	0 °C	-6 °C	-3 °C
12/19/2006	84	-9.7 °C	-14 °C	-4 °C	-7 °C	-7 °C	-14 °C	-10 °C
12/21/2006	64	-6.7 °C	-9 °C	-6 °C	-3 °C	-3 °C	-6 °C	-4 °C
12/22/2006	78	-3.4 °C	-6 °C	-1 °C	2 °C	2 °C	-11 °C	-4 °C

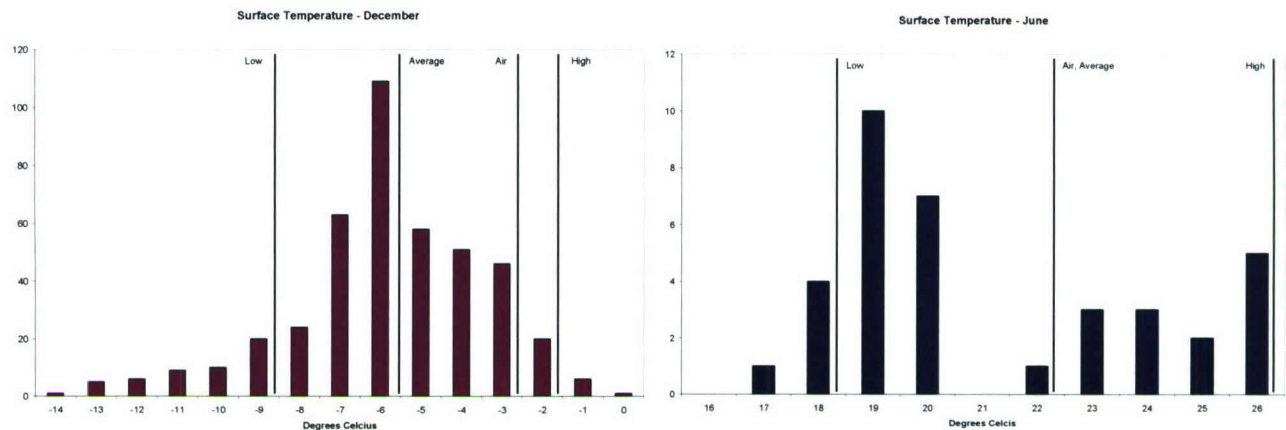


Figure 2. Frequency Histogram of Waste Surface Temperatures

### *Compacted Unit Weight*

Preliminary tests were conducted on the active face of the landfill under normal operation. Records were analyzed for normal operations of waste placement. Water addition was also used on the active face. Steps of the process included spraying the waste with water from a water truck, compacting at the modified water contents (Figure 3), and analyzing the resulting topographic and incoming waste data from the GPS system together with site records. Difficulties were encountered in this method due to inconsistencies in rate of waste delivery during the day (in comparison to the application of moisture), the limit of rate of application of water using a water truck (primarily due to cycle time of the machinery, i.e., water filling periods), and difficult terrain for access for the water truck when the site activity was at steep corners of the site.



Figure 3. Water addition to waste mass using turret nozzle and compaction under wet conditions

During a period of technical problems with the GPS system, an attempt was made to determine the in-place compacted density of the waste using a localized sampling method. A

sample of compacted waste was removed from the active face of the landfill and placed in a roll off truck. The weight of the sample and the volume of the hole were used to determine the unit weight. A backhoe removed a sample of waste and placed it into the can of an empty roll top truck (weighed prior to placing sample in can). The outline of the hole left by the backhoe bucket was then marked with spray paint and a PVC Grid was placed over the hole (Figure 4). The PVC Grid was constructed of 38 mm (1.5 inch) diameter PVC pipe. Each intersection in the grid was secured with at least one screw, two cable ties, and electrical tape. The screws and cable ties were used to provide rigidity for the grid, and the electrical tape was used in order to cover the sharp edges of trimmed cable ties. The grid was 1.5 m by 1.5 m with five rows spaced at 0.3 m and ten columns spaced at 0.15 m. The grid was used in order to aid in the determination of the volume of the hole left by the backhoe bucket. Depth measurements were taken at specific points on the grid using a retractable tape measure. The volume of the hole was determined by summing the products of the average depth for each cell within the grid and the area of the corresponding cell. The localized compacted waste unit weight values obtained from test samples from the active face are presented in Table 3. The results of this localized unit weight determination indicate the problems of scale associated with such experiments. The calculated density values are between 50 and 125 times smaller than the in-place waste and cover unit weight of  $11.8 \text{ kN/m}^3$  reported by the landfill for 2005. This reinforces the need for larger scale testing for determination of compacted unit weights.



Figure 4. Localized compacted unit weight determination

Table 3. Summary of Local Compacted Waste Unit Weight

	Weight (kN)	Volume ( $\text{m}^3$ )	Unit Weight ( $\text{kN/m}^3$ )
Sample 1	0.98	10.25	0.095
Sample 2	1.42	6.80	0.210

#### *Sensitivity to Temperature*

The compacted unit weight was monitored during normal operations at the landfill to address seasonal temperature variation on compaction characteristics. A plot of all measured unit weights vs. average daily temperature is presented in Figure 5. No discernable trend with temperature is observed for normal compaction operations. This represents normal operations

and not the highly controlled and smaller scale test plots that were conducted to address influence of moisture addition.

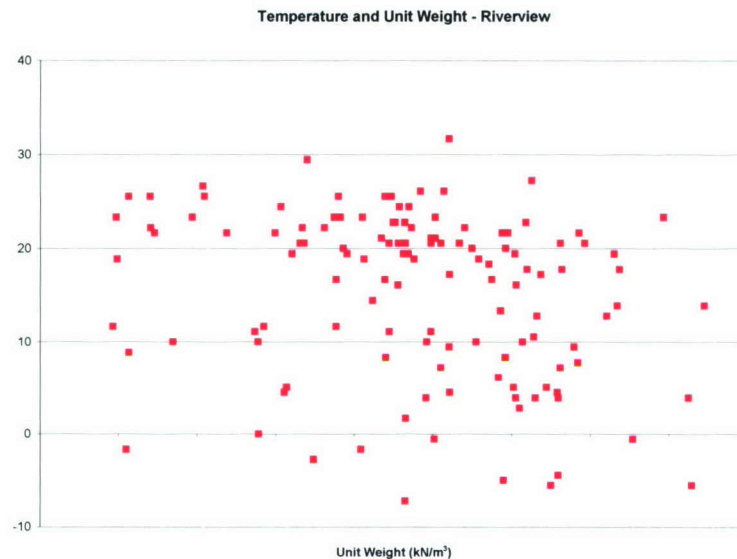


Figure 5. Relationship between Compacted Unit Weight and Moisture Content

#### *Modification of Compaction Practices at Controlled Test Plot*

Field compaction tests in the study involved compaction at prescribed conditions (compactive effort, moisture content, type of waste, and temperature) in the designated test plot. In addition, normal waste compaction practices were monitored over the duration of the project to evaluate global variability in resulting compacted unit weights. Wastes were compacted by Riverview personnel in accordance with the requirements of the test program. Data collected for the compacted wastes consisted of weight, volume, and moisture content measurements. The GPS setup was used to precisely monitor compaction conditions at the site. Exact volumes were calculated using the GPS system. Waste weights were determined using measurements obtained at the scale-house of the facility. The weights and volumes were used to determine as-placed unit weights of wastes. Each test was conducted on approximately 40 t of waste. Specifications for the tests are presented in Table 4. Photographs of the test plot are presented in Figure 6.

Table 4. Controlled Test Plots

Date	Units	11-Dec-06	23-Jan-07	27-Dec-06	15-Feb-07	13-Dec-06	12-Dec-06	19-Dec-06
Waste Quantity	kN	377.3	334.7	355.3	249.7	296.9	296.9	361.6
Air Temperature	°C	10	-3.9	1.7	-5.6	7.8	7.8	7.8
Water Added	Liter	0	3780	7560	7560	11340	11340	15120
Water Temperature	°C		13.3	12.2	12.2	11.7		-17.8
Volume	m <sup>3</sup>	68.0	41.1	92.9	37.0	27.7	27.7	36.1
Moist Unit Weight	kN/m <sup>3</sup>	5.55	8.15	3.82	6.75	10.72	10.72	10.01

Moisture Content	%	28%	42%	54%	66%	76%	76%	80%
Dry Unit Weight	kN/m <sup>3</sup>	4.34	6.38	2.99	5.28	8.39	8.39	7.83

Results of the test plots are presented in Figure 7. A significant increase in dry unit weight was obtained for moisture contents in the 40-80% range. A strong relationship existed between moisture content and compacted dry unit weight. Optimal moisture content of approximately 65% resulted in a maximum dry unit weight of approximately 8.5 kN/m<sup>3</sup>. Much lower dry unit weights (4 kN/m<sup>3</sup>) resulted for both dry of optimum and wet of optimum conditions. It is observed that nearly twice the waste can be placed within a given volume with a moderate addition of water during compaction.



Figure 6. Test Procedures for Controlled Test Plot

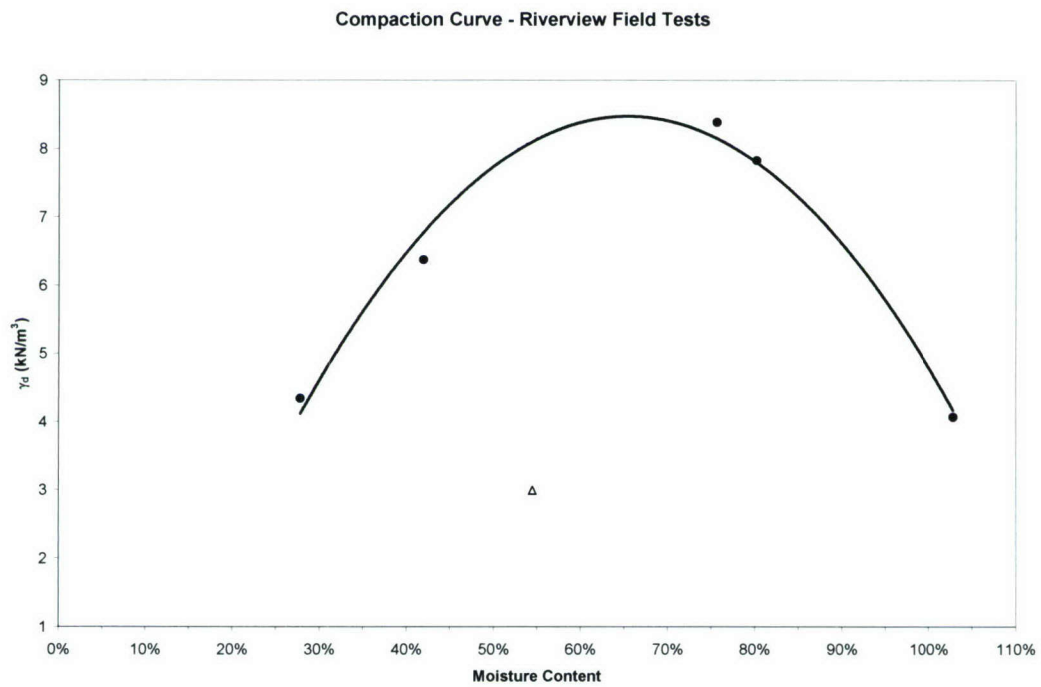


Figure 7. Results from Compaction Test Plots

On Figure 7, the datapoint associated with 55% moisture content has not been included in the curve fit. This datapoint represents a field compaction test that was conducted on the first local curbside waste collection after the Christmas holiday season. It is understandable that a fundamentally skewed waste stream was present for this test due to the large amount of packaging and gift wrap as compared to the conventional municipal waste stream. The partner landfill has agreed to repeat a test at this moisture content to obtain more representative test data. Due to freezing weather conditions at the site, this test has not yet been able to be completed (as of February 15, 2007).

#### *Baseline Laboratory Compaction Tests*

A series of compaction tests was conducted on manufactured municipal solid waste using a 150 mm diameter compaction mold. The waste fractions were based on EPA guidelines for average disposal trends in 2003 (EPA 2003). The tests were conducted in accordance with ASTM D-1557 (Modified Proctor Compaction Effort). Seven tests were conducted at gravimetric moisture contents ranging from 10% to 150%. The compacted unit weight was sensitive to compaction water content with an optimum water content (that which provided maximum dry unit weight) of approximately 50% (Figure 8). The maximum dry unit weight was determined to be 5.6 kN/m<sup>3</sup>. The results of waste compaction are consistent with general trends observed for fine grained soils, notably presence of an optimum moisture content for compaction.

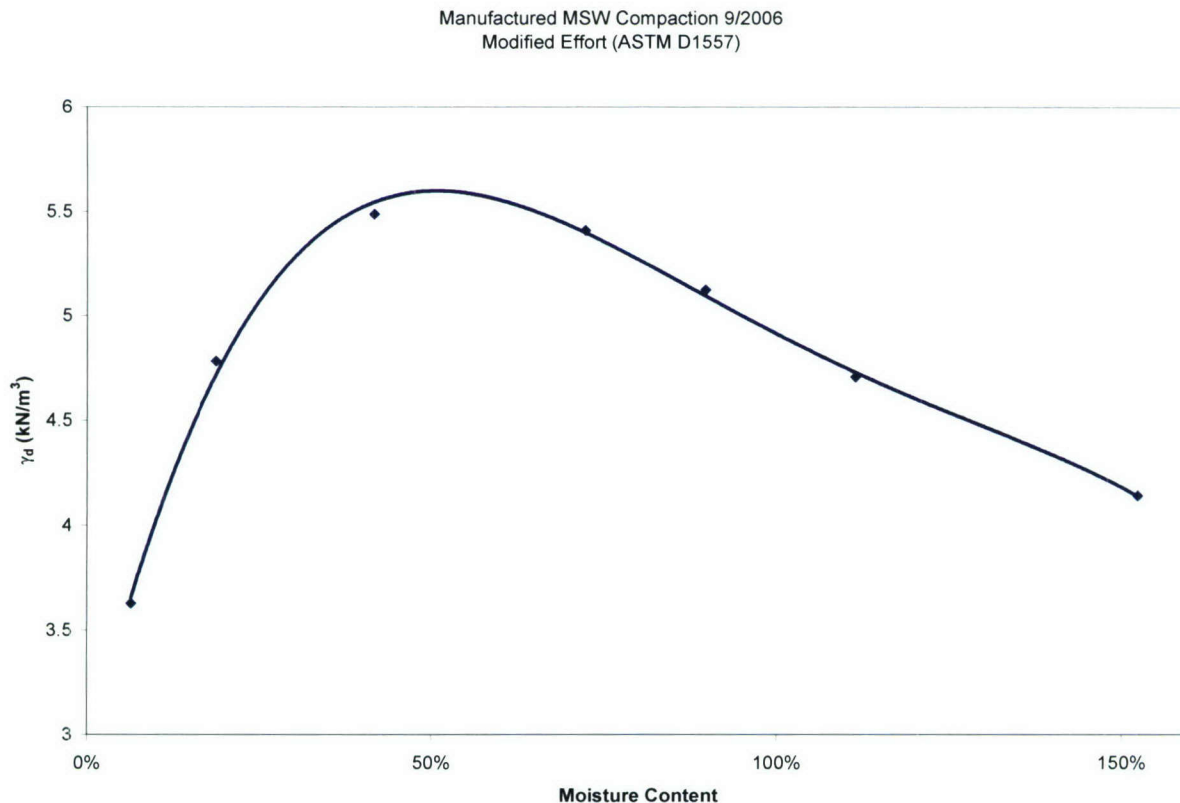


Figure 8. Results of baseline laboratory compaction tests

### *Waste Compressibility as a Function of Moisture Content*

Further laboratory tests were conducted to determine the influence of moisture content on the compressibility of compacted wastes. Relatively large-scale specimens were prepared to match the compacted unit weights of the baseline laboratory experiments. A cylindrical mold (300 mm diameter, 350 mm height) was used for this test program. Samples were compacted using static load application from a 700 kN-capacity loading frame. Samples were prepared using 7 lifts of waste within each mold. Upon reaching designated moist unit weight, the specimens were subjected to compression tests at a rate of 2.5 mm/min. Constituents of manufactured wastes were mixed in ratios consistent with data from the EPA (2005). The test setup for compression tests and representative waste constituents used for a given compression test are presented in Figure 9.

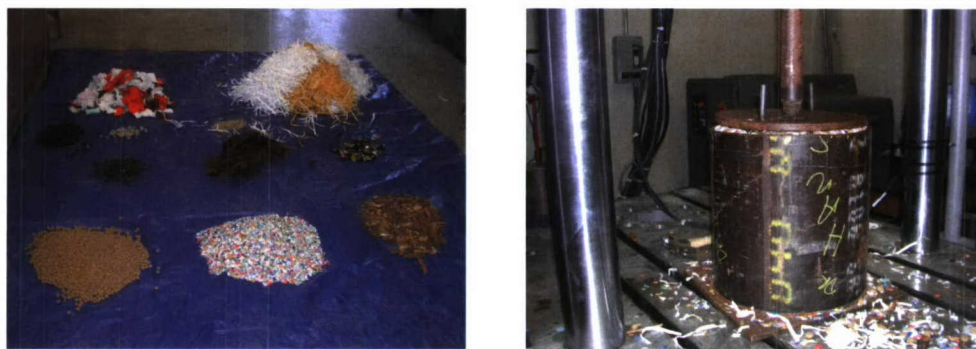


Figure 9. Constituents of manufactured waste and compression test

Data from the compression tests are presented in Figure 10. The data is plotted as axial strain vs. log confining stress (a typical format for presenting soil compression data). The slope of this line is termed the compression ratio. Relative stiffness of the wastes can be determined by comparing compression ratios. Compression ratio varied from approximately 0.17 to 0.21 over the range of moisture conditions tested. Secant modulus at 25% strain was also determined from the tests. Secant modulus ranged from approximately 1100 to 4300 kPa. Both the compression ratio and the secant modulus are plotted against moisture content in Figures 11 and 12. Both generally follow a similar trend of increasing stiffness at optimal moisture conditions. It was determined that stiffness (as quantified using either compression ratio or secant modulus) varied as a function of moisture content in a relationship similar to unit weight (generally a bell-shaped curve). A plot of secant modulus vs. dry unit weight is presented in Figure 13. It is observed that secant modulus is generally correlated to dry unit weight.

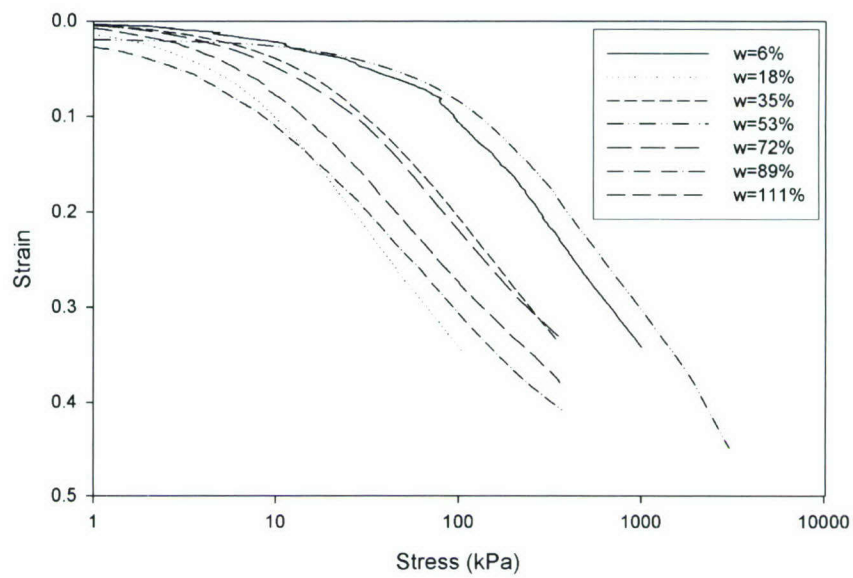


Figure 10. Results of Laboratory Compression Tests on Wastes

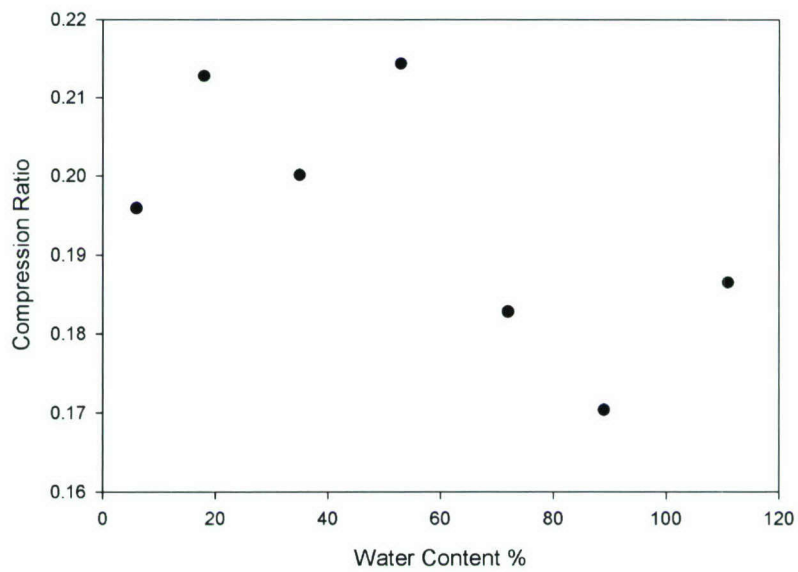


Figure 11. Compression Ratio vs. Moisture Content

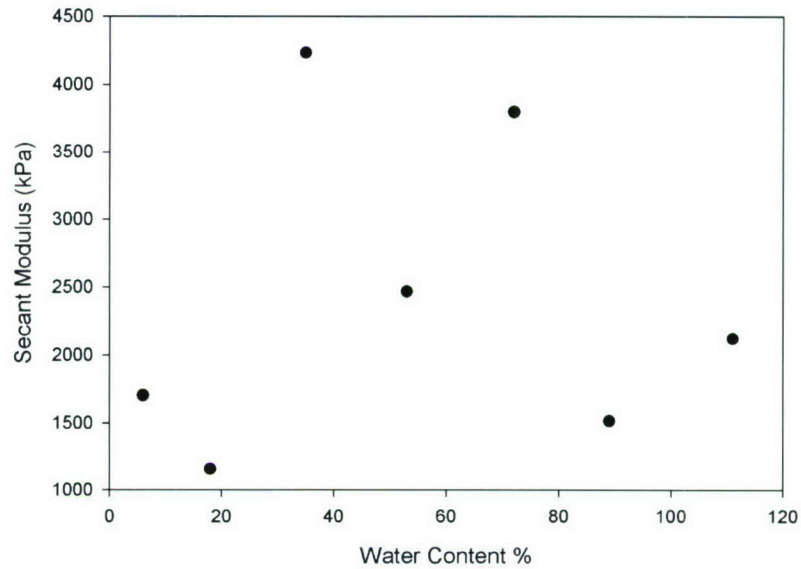


Figure 12. Secant Modulus vs. Moisture Content

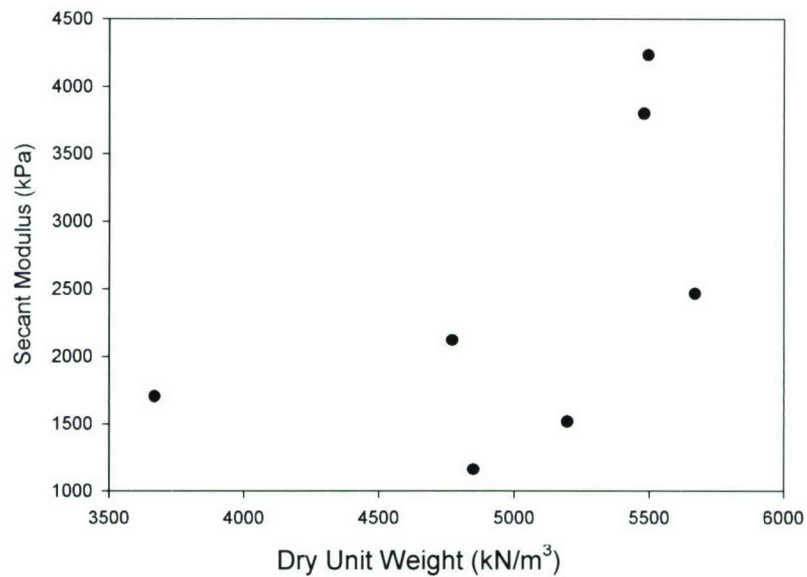


Figure 13. Secant Modulus vs. Dry Unit Weight

### Navy Wastes

Extension of the results of this investigation to specific waste streams (such as wastes generated by the Navy) is possible and the structural composition of such waste streams is generally consistent with that of MSW landfills. The results of this project are of high relevance to ONR/DOD due to the large magnitudes of wastes (municipal-grade as well as hazardous-

grade wastes) produced by the military. Land-borne waste production is significant at military bases across the world. Some waste streams within the military structure (e.g., munitions) represent “legacy wastes” that can represent particularly long-term environmental liabilities. Large magnitudes of waste are also produced by naval operations. While disposal at sea is currently in wide use for the U.S. Navy (EPA 2006), it is expected that more landfilling will be required in the future due to increased global environmental pressures. The results from this project can be applied for cost savings of landfilling as well as environmental conservation.

### **Summary and Conclusions**

Results of the field experiments indicate that compaction of wastes is highly variable during normal operation at a landfill. Measured compacted dry unit weights ranged from 0.9 to 8.0 kN/m<sup>3</sup> (average 4.8 kN/m<sup>3</sup>). No relationship between compacted dry unit weight and temperature was observed during evaluation of normal operations at the landfill. A strong relationship existed between moisture content and compacted dry unit weight. Optimal moisture content of approximately 65% resulted in a maximum dry unit weight of approximately 8.5 kN/m<sup>3</sup>. Much lower dry unit weights (4 kN/m<sup>3</sup>) resulted for both dry of optimum and wet of optimum conditions. Results from laboratory tests indicated a similar fundamental trend and lower compressibility of wastes was observed for optimum compaction conditions. Significant cost savings and environmental conservation can be achieved by manipulating moisture content during waste compaction.

This study provides significant improvements in the state-of-the-art and the state-of-the-practice of waste containment and management in the U.S. Results from the proposed study provide both environmental and economic benefit to a wide range of constituents including landfill host communities, landfill owners and operators, government agencies, and the general public. Long-term cost savings can be achieved due to reduced amount of airspace required for waste disposal. Using the test site as an example, a savings of approximately \$4.6 Million would be realized if compaction efficiency was increased so that 30% more waste was placed in each m<sup>3</sup> of remaining airspace. The test results from the test plots indicated a potential 100% increase in compaction efficiency was possible. A lower number is used for these calculations to provide a conservative estimate. By placing more waste into the footprint of the cell, environmental benefits are also realized in the reduced need for additional landfills.

The implementation of the GPS system provides more representative measurements of geometric attributes of the compaction test cells than the approximate measurements that have been available in the past. Overall, the study has impact for landfills located in the entire U.S. and will lead to further studies of effects of compaction conditions on engineering properties of wastes. Optimum waste compaction is a fundamental requirement for sustainable environmental management of landfills in the 21<sup>st</sup> century. A significant outcome of the project is a database that provides an ability to develop formal guidelines for appropriate compaction methods to achieve target design values for engineering properties of wastes. The guidelines for compaction will be broadly applicable to conventional landfills as well as to sites that employ alternative landfilling techniques (e.g., bioreactor landfills) or accept unique waste streams (such as military or other wastes).

## References

- Bobeck, R. (2004). Personal Communication. City of Riverview Solid Waste Division, Riverview, Michigan.
- Bleiker, D. E., Farquhar, G., and McBean, E. (1995). "Landfill Settlement and the Impact on Site Capacity and Refuse Hydraulic Conductivity." *Waste Management & Research*, 13, 533-554.
- Briaud, J. L. (2005). "Intelligent Compaction Systems: Overview," *Transportation Research Board 84<sup>th</sup> Annual Meeting, Compendium of Papers CD-ROM*, National Academies, Washington, D.C.
- CIWMB (2005). Landfill and Other Solid Waste Facilities, Solid Waste Disposal Tonnage Summary Data, California Integrated Waste Management Board, <http://www.ciwmb.ca.gov/Landfills/LFData.htm>, last accessed July 28, 2005.
- EPA (2005). *Municipal Solid Waste Generation, Recycling, and Disposal in the United States: Facts and Figures for 2003*, EPA530-F-05-003, United States Environmental Protection Agency, Office of Solid Waste and Emergency Response.
- EPA (2006). Federal register environmental documents, Notice of Record of Decision for the Disposal of U.S. Navy Shipboard Solid Waste from Surface Ships, <http://www.epa.gov/fedrgstr/EPA-IMPACT/1997/February/Day-14/i3783.htm>, last accessed October 29, 2006.
- Gabr, M. A., and Valero, S. N. (1995). "Geotechnical Properties of Municipal Solid Waste." *Geotechnical Testing Journal*, ASTM, 18(2), 241-251.
- Holtz, R. D. and Kovacs, W. D. (1981). *An Introduction to Geotechnical Engineering*, Prentice Hall, Upper Saddle River, NJ.
- PIER (2005). "Landfill Gas to Energy," Public Interest Energy Research, California Energy Commission, <http://www.energy.ca.gov/pier/renewable/biomass/landfill/>, last accessed July 24, 2005.
- Qian, X. Koerner, R. M., and Gray, D. H. (2002). *Geotechnical Aspects of Landfill Design and Construction*, Prentice Hall, Upper Saddle River, NJ.
- Sharma, H. D. and Lewis, S. P. (1994). *Waste Containment Systems, Waste Stabilization, and Landfills – Design and Evaluation*. John Wiley and Sons, Inc., NY.
- Tchobanoglous, G., Theisen, H. and Vigil, S. A. (1993). *Integrated Solid Waste Management: Engineering Principles and Management Issues*, McGraw Hill, New York, 978 pp.

## Publications and Presentations Related to Project

### Abstracts

*Hanson, James; 'Studies of Properties and Response of Waste Containment Systems'; Presentation to California Integrated Waste Management Board, June 26, 2006; Presentation to Los Angeles County Sanitation District, August 29, 2006; Presentation to Monterey Regional Waste Management District, September 29, 2006; Panel Discussion at American Society of Civil Engineers GeoDenver 2007, February 20, 2007.*

This study was conducted to develop procedures for optimization of waste compaction at Municipal Solid Waste (MSW) Landfills. Optimization using GPS technology can improve waste containment efficiency, resulting in conservation of land and environmental resources. A fundamental understanding of waste compaction processes or widely acceptable guidelines for industry use are currently not available. A field study was conducted to identify the baseline parameters and conditions that control compaction of wastes. Field compaction analyses were

conducted in a test plot that was part of an active landfill cell at Riverview Land Preserve (a municipal government operated, MSW landfill in Riverview, Michigan). The field study was conducted at a scale appropriate to the waste management industry with sample sizes of approximately 355 kN. A GPS system was used to make the required geometric calculations for determination of in field densities and placement efficiency of wastes in a manner that has not been previously possible. Results indicate that compaction of wastes is highly variable during normal operation at a landfill. Measured compacted dry unit weights ranged from 0.9 to 8.0+ kN/m<sup>3</sup> (average 4.8 kN/m<sup>3</sup>). A strong relationship existed between moisture content and compacted dry unit weight. Optimal moisture content of approximately 65% resulted in a maximum dry unit weight of approximately 8.5 kN/m<sup>3</sup>. Much lower dry unit weights (4 kN/m<sup>3</sup>) resulted for both dry and wet of optimum conditions. Significant cost savings and environmental conservation can be achieved by manipulating moisture content during waste compaction.

**Production, characterization and self-assembly of passivated and  
functionalized metal nanoparticles  
and  
collaborative use of Zeiss-10 transmission electron microscope**

Project Investigator:

Steven A. Harfenist  
Physics Department  
California Polytechnic State University  
San Luis Obispo, California

## Introduction:

The goals of this first year of funded research were on the most part met. First was to successfully install the nanocrystal flow processor (NXFP). Next we were to start characterizing the NXFP for effects of temperature and gas flow rates on nanocrystal size and shape. One goal was to size-select aliquots of nanoparticles from one production run. Another goal was to try to produce nanoparticles not previously produced using this method. And last were to successfully train students on and collaboratively use the newly acquired Zeiss-10 Transmission Electron Microscope. Out of these five goals, four were successfully accomplished.

## Results:

The NXFP was setup mainly by the PI (Harfenist) during the spring quarter of 2006. This was a good use for my release time from teaching during that quarter. Gas lines, regulators, flow meters thermocouples and data logger and a high voltage source were purchased and/or acquired through the Physics Department and successfully installed.

Two majors from the Physics Department were employed mainly during the summer of 2006 to do the nanoparticle production. Approximately eight successful production runs were performed with some variance in processing conditions. It was this part, the careful characterization of the NXFP through variation in processing conditions that was unsuccessful. We discovered late in the summer that the thermocouple feed-through used

to monitor the crucible temperature were particular to a specific type of thermocouple.

Therefore our temperature readings were inaccurate and characterization of the processor with crucible temperature was not possible. This problem has been ameliorated; we purchased the correct, C-type, thermocouple and extension wire that should give us more accurate temperature readings. The students

did successfully produce silver nanoparticles passivated with dodecane- and hexane-thiol self-assembled-monolayers. An example of the hexanethiol passivated nanoparticles is shown in figures 1 and 2. Here the nanoparticles were imaged with the newly acquired Zeiss-10 Transmission Electron Microscope (TEM) and have an average diameter of 4.3nm. The nanoparticles were collected on a liquid nitrogen-cooled cold finger and then sonicated in toluene to disperse the nanoparticles.

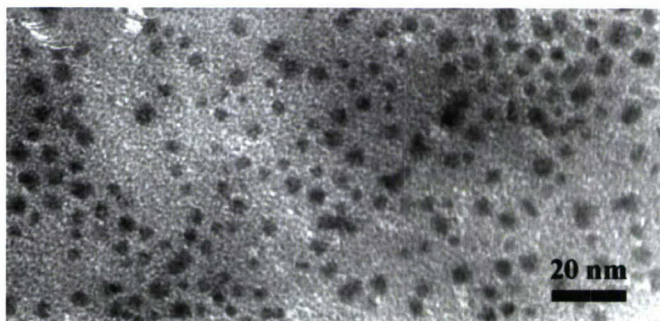


Figure 1. Hexanthiol passivated Ag nanoparticles of average diameter 4.3nm. Magnification = 200Kx.

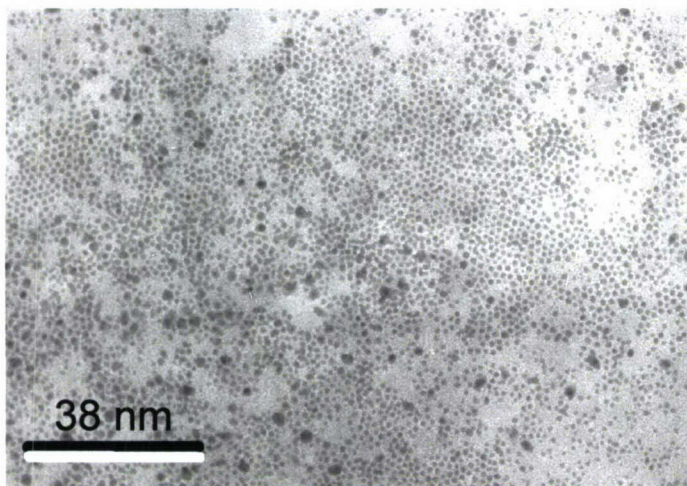


Figure 2. Transmission Electron Microscope image of hexanethiol coated Ag nanoparticles with average diameter of 4.3nm. Magnification = 100kx.

acetone and collecting the resultant precipitate. Theoretically, the last solution should have the smallest-sized particles, while the initial solutions should have the largest.

We had very limited success with this method. Shown in figure 3 are size distributions of nanoparticles that were to be size-selected. We only considered the precipitates formed, expecting them to have the desired sizes. Here the samples B, D, F and H should have decreasing size from B through F. The raw solution was first diluted by 50% with acetone. The precipitate was then separated and dispersed in toluene; the first to be processed like this is sample B. Then the remaining, unprecipitated solution was further diluted by 50% with acetone and the remaining precipitate was dissolved in toluene as sample D. This process continued until sample F and H were produced. The average particle sizes did not follow the trend expected; the average size should decrease from B to H. What is evident is that the initial distributions

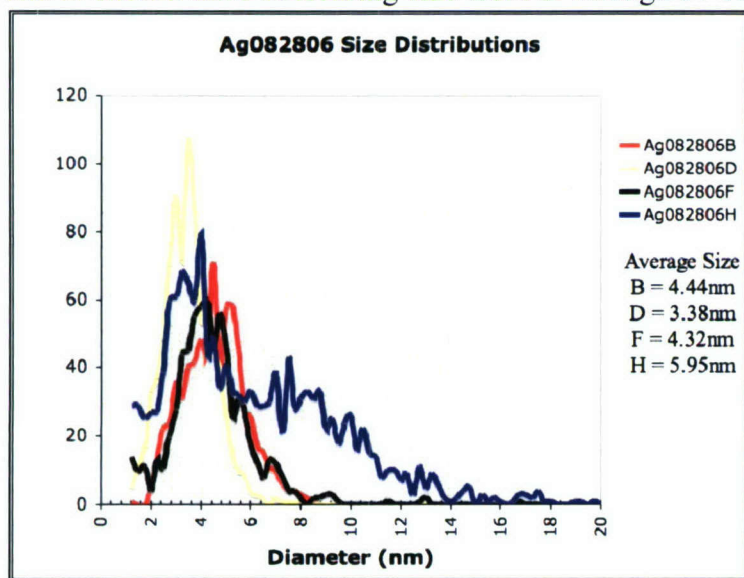


Figure 3. Size distributions of Ag nanoparticles size-selected from a single solution. Diameters are indicated assuming the particles are spherical.

became more peaked, while the later solutions were broader. A number of problems that might have caused this result were centrifuging the precipitate rather than letting it naturally settle and using acetone instead of other, less volatile, solvents. The first problem emphasizes a general problem associated with "dissolved" nanoparticles. These

Since we had inaccurate temperature data, we decided to explore the method of separating out different sized aliquots of nanoparticles from a single solution. This method is called size-selective-precipitation. The idea is that when one adds a solvent that the nanoparticles are not soluble in, for example, acetone. The largest particles will start to coagulate, flocculate and come out of solution. The precipitate is then collected and the collected, larger, particles are then redissolved. This process continues with adding more

by 50% with acetone. The precipitate was then separated and dispersed in toluene; the first to be processed like this is sample B. Then the remaining, unprecipitated solution was further diluted by 50% with acetone and the remaining precipitate was dissolved in toluene as sample D. This process continued until sample F and H were produced. The average particle sizes did not follow the trend expected; the average size should decrease from B to H. What is evident is that the initial distributions

materials are not truly dissolved in terms of the chemical definition. Therefore, centrifuging could cause many, non-precipitated, nanoparticles to be collected, rather than only those desired. Acetone is a very volatile solvent and its high vapor pressure may not work well with this method of size-selection.

Optical absorption spectra of the nanoparticle solutions were obtained on an Optical Absorption Spectrophotometer in Dr. Chad Imuus' laboratory.

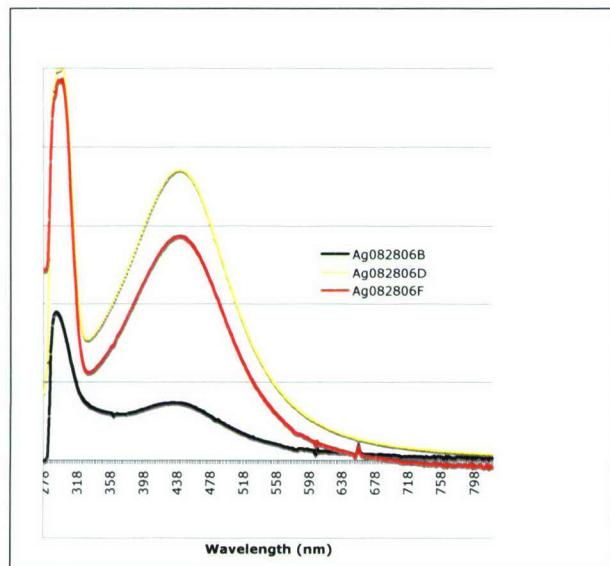


Figure 4. Optical absorption spectra of Ag nanoparticles.

wavelength decreases with decreasing size of the nanoparticle. Shown in figure 4 is the optical absorption spectrum of three nanoparticle solutions. The three spectra shown here have absorption wavelengths varying from 434nm to 438nm. The variation does not agree with the size distributions given in the previous figure. The large width of the size distributions is probably why the absorption spectra do not shift as expected. This is proof that the particles are made of Ag. More careful size selection should give better results for absorption spectra shifts versus nanoparticle size.

A new type of nanoparticle was produced using the NXFP aerosol method. Here, indium nanoparticles were produced and were tentatively passivated with trioctylphosphine oxide (TOPO). The result is tentative about the passivation because chemical analysis of the passivation has yet to be done. The particles were verified as being In from the large angle, electron diffraction pattern (not shown since contrast was poor). Figure 5 is a TEM image of In nanoparticles collected directly on the TEM

Absorption Spectrophotometer in Dr. Chad Imuus' laboratory. A metal nanoparticle will experience collective electron excitation when in an external field. When the frequency of the external field is near the resonance frequency of this collective excitation (the plasmon resonance) strong absorption will occur. For silver nanoparticles the plasmon resonance is at approximately 440nm wavelength. Since this is near the blue part of the optical spectrum, the transmitted light is peaked near the orange or yellow wavelengths. The plasmon resonance wavelength for nanoparticles is dependent upon the nanoparticle size and shape and can be used as a measure of average nanoparticle size where the resonance

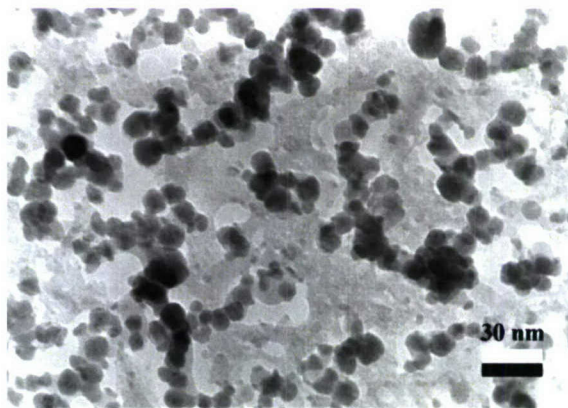


Figure 5. Indium nanoparticles produced in the NXFP. Average size is 10nm diameter. Scale bar is 30nm.

grid from the NXFP's exhaust output. The average diameter of the particles was 10.0 nm. The distribution was quite wide, with a standard deviation of 3.9nm. Collecting the samples directly onto the TEM sample grid enabled us to determine indium nanoparticle production rates.

There were approximately 610 nanoparticles in the original image that spanned 425nm x 350nm. This was collected over a 30s time span. Given that the NXFP has an exhaust port size of 1.5cm in diameter, we can conclude that for this particular gas flow rate and temperature, that approximately  $3.66 \times 10^{10}$  indium nanoparticles of 10.0nm diameter were produced every second. Using an image processing program (ImageJ) I was able to calculate the total volume of the nanoparticles in the image which and from that an average indium mass flow rate (evaporation rate) of  $2.14 \times 10^{-7}$  g/s. Here we presumed a 100% sticking probability for the nanoparticles on the TEM grid. The nanoparticles were warm and the grid cold, thus it is a safe assumption from thermodynamic principals that most of the particles stuck to the TEM grid.

Both of my research students were trained on the Zeiss TEM and they took many of the images shown here. Other students were also trained on the electron microscope, a Master's student in Dairy Sciences and another in Chemistry. Details of their progress follow immediately.

#### Collaborative use of Zeiss-10 TEM:

In addition to having trained my research students from this past year on the TEM, two other groups have formed collaborations with me for using the TEM in their research. Dr. Ray Fernando is studying coatings that have aluminum oxide nanoparticles as a constituent. Dr. Fernando's student, Margarita Gonzalez, was trained in all aspects of the TEM, including dark room developing techniques. Shown in figures 6 and 7 are TEM images of the alumina nanoparticles from their samples. Here the average size is approximately 20nm, though a more thorough statistical analysis is forthcoming. Figure 7 shows a much higher concentration at 200kx

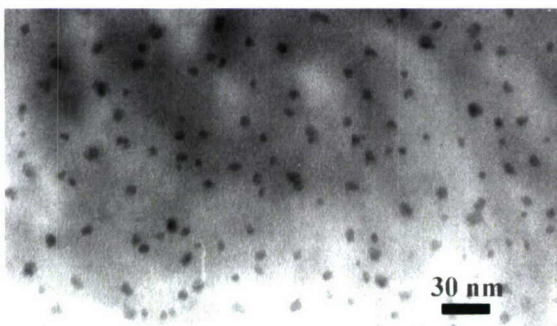


Figure 6. TEM image of alumina nanoparticles. Avg. Size ~20nm.

magnification as opposed to the relatively lower concentration in figure 6 at 160kx magnification. Another interesting use of the TEM is for imaging biological samples. Jessica Lee, a Master's student in Dr. Jimenez' group in Dairy Sciences is studying the vesicle structures formed from liposomes. She found an interesting paper on how to negatively stain the liposomes and the results are shown in figure 8. What is seen here are multi-lamellar structures formed from soy liposomes. It is believed that the bright contrast regions are the liposomes, while the darker are the regions around them.

Imaging these biological materials is new for me and a

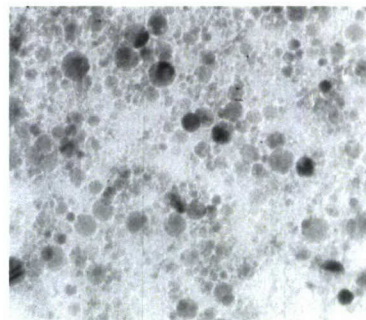


Figure 7. Transmission Electron Microscope image of  $\text{Al}_2\text{O}_3$  nanoparticles.

considerable amount of time went into perfecting the imaging process. On the whole, as can be seen in most all of the images, there is still a problem with getting even contrast over the entire image. I have ruled out a number of reasons why this might be and am left with either the safe lights used in the dark room are not the right color and effect the unprocessed film, or the microscope is somehow slightly exposing the film before or after it is due to be exposed normally.

All of the research shown here will be continued throughout this coming year. Scott Lewis, one of my original undergraduate research students has finished his senior project on the production and characterization of nanoparticles and is graduating this year. I expect to recruit at least one more student for this summer and continue to work with my other research student, Steve Tomak, another physics major.

With the necessary materials to accurately determine the NXFP crucible temperature, we should be able to finish the goal of characterizing the nanoparticles produced with changes in temperature and gas flow rates. Further characterization of indium nanoparticles will continue with some samples being sent out for chemical analysis.

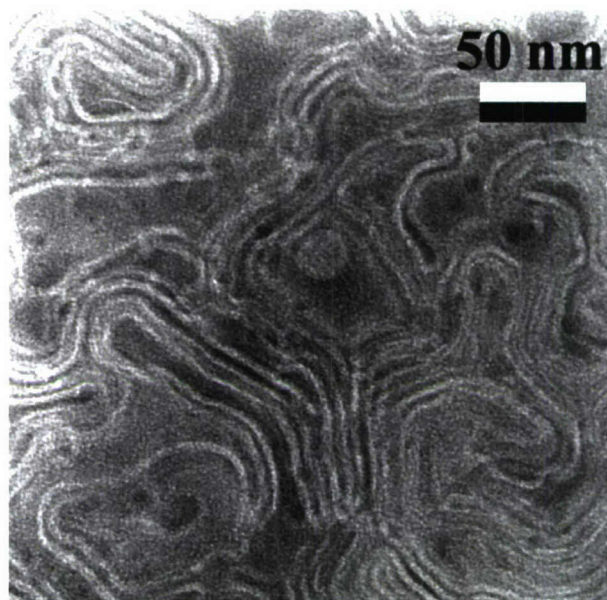


Figure 8. Multi-lamellar structures formed by soy liposomes. Scale bar is 50nm.

## **Development of reactive polymer coatings for chemical weapons degradation**

Project Investigator:

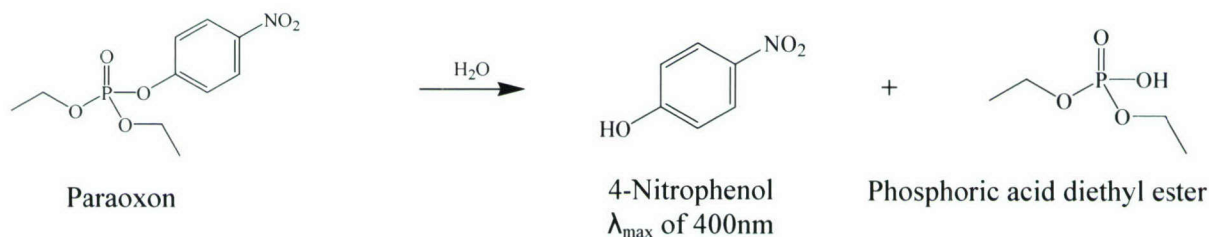
Chad Immoos  
Chemistry and Biochemistry  
California Polytechnic State University  
San Luis Obispo, California

## Abstract:

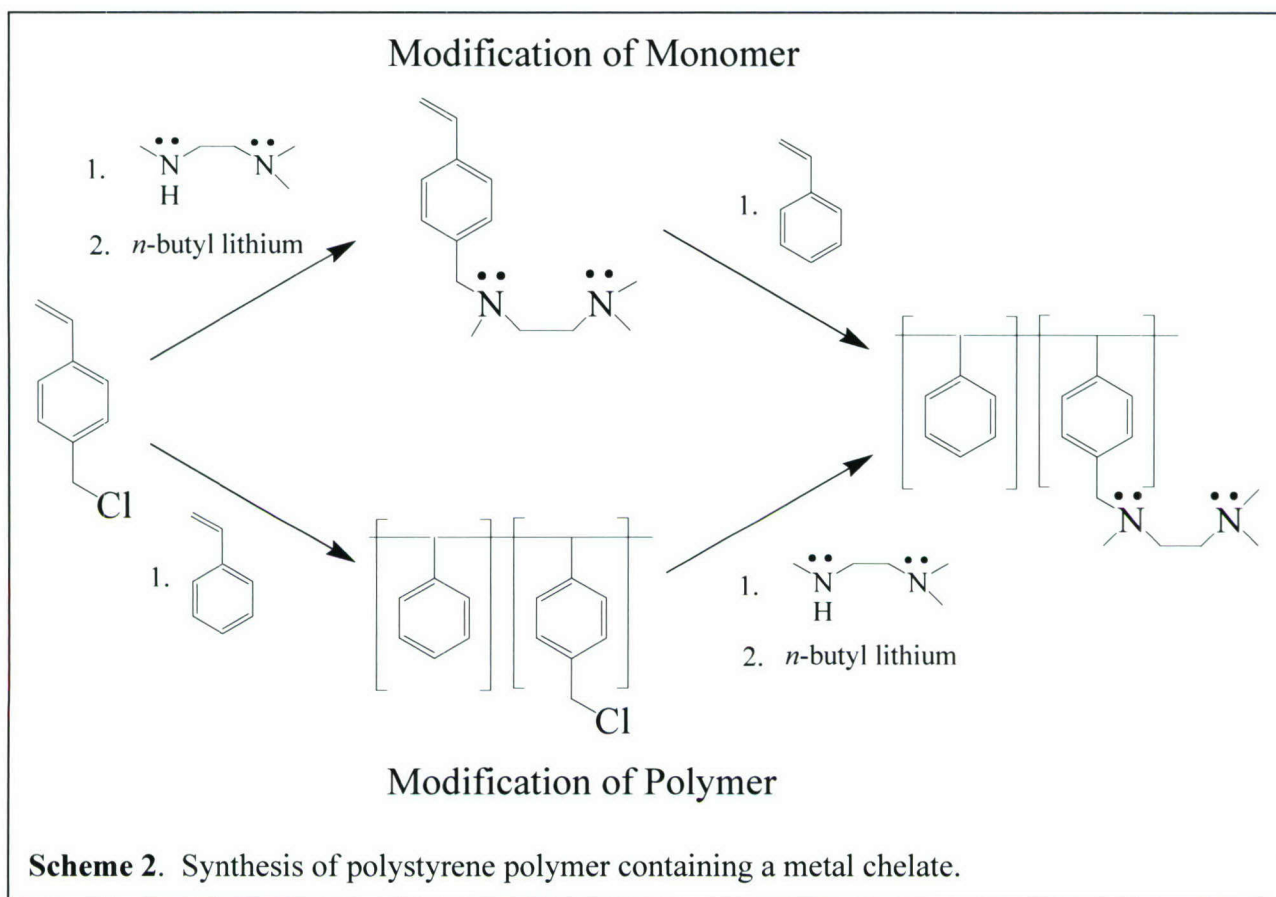
Chemical warfare agents are extremely hazardous materials and are a direct threat to military personnel assigned to interact with these compounds. Among these compounds are certain organophosphate, phosphinate, and phosphonate esters, such as VX, Sarin, and Soman.<sup>1</sup> As such, the development of effective and rapid methods for the degradation of these compounds is of great interest to the Department of Defense. The objective of this proposal is to develop new technology for the degradation of chemical warfare agents by catalysts immobilized in polymer matrixes. We will synthesize polymers containing transition metal catalysts or reactive metal nanoparticles in an effort to develop polymer coatings capable of degrading chemical warfare agents under ambient conditions. Specifically, we will immobilize transition metal catalysts or reactive nanoparticles previously reported to have enhanced reactivity toward CW agents and monitor the degradation of CW agents by the polymer composites.

## Preliminary/Previous Results

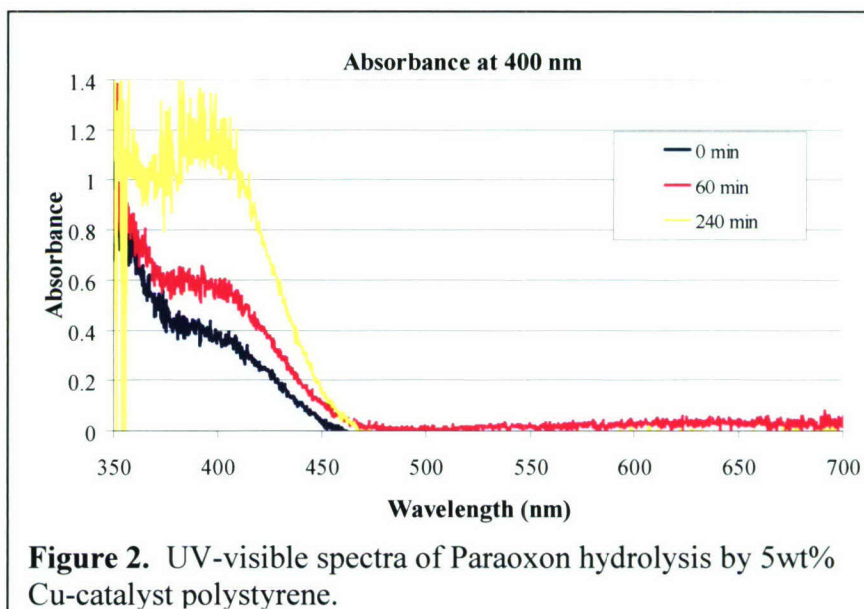
*Development of transition metal catalyst-containing polymers for chemical warfare agent degradation.* We have synthesized polymers containing copper complexes and screened them for reactivity toward Paraaxon, a chemical warfare agent simulate (Figure 1). The hydrolysis of Paraaxon, as shown in Scheme 1, was monitored via UV-visible spectrophotometry.<sup>15</sup> The reactivity of several copper complexes toward Paraaxon and the kinetics of the hydrolysis reaction were analyzed.



**Scheme 1:** Hydrolysis of Paraaxon to generate 4-nitrophenol.

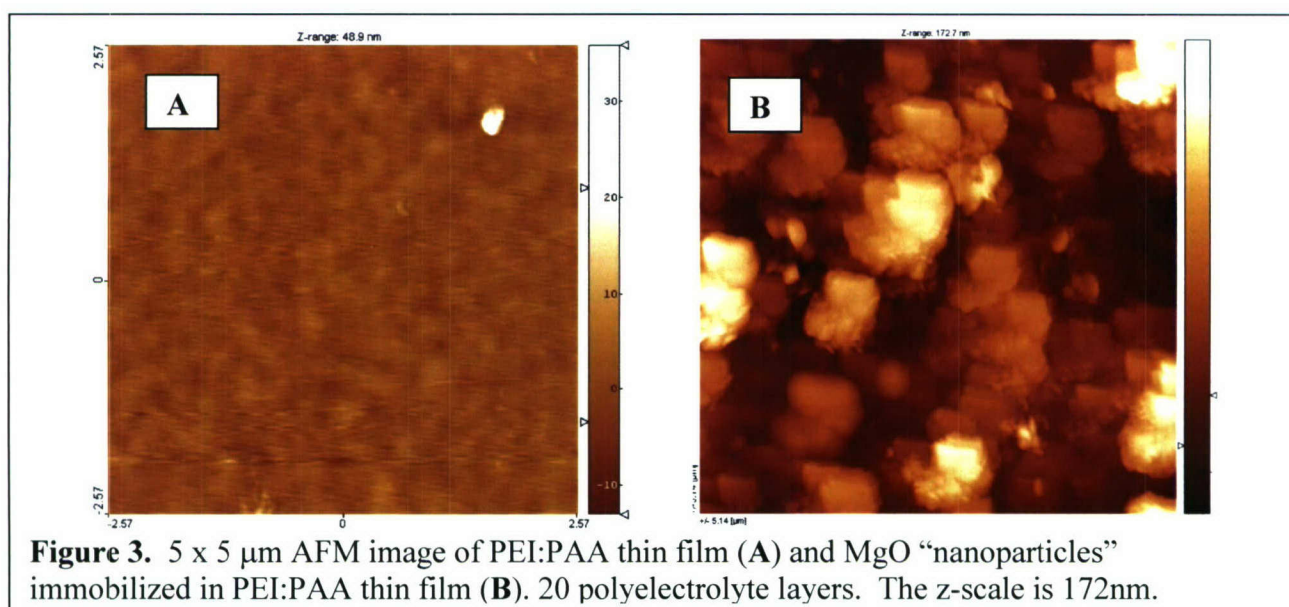


A copper-loaded styrene polymer was successfully synthesized according Scheme 2. Modification of the monomer was explored as a possible route to the generation of chelate-containing polymers. If successful, we intended to use this monomer to make both linear polymers and polymer nanoparticles. Synthesis of the monomer, although difficult because of solubility issues was successful, however, it was found that the amines inhibited the free radical polymerizations. Attempts to overcome this obstacle were unsuccessful. Alternatively, modification of the styrene polymer with trimethylethylenediamine was successful and produced a copper-loaded styrene polymer. Styrene can be co-polymerized with varying quantities of chloromethyl styrene to generate the desired copolymer. Addition of the trimethylethylenediamine chelate can be monitored by loss of the C-Cl bond via FTIR analysis. The



synthesized polymer was shown by UV-vis spectrophotometry to be effective at hydrolyzing Paraoxon.

*Development of reactive nanoparticle-containing polymer films for chemical warfare agent degradation.* We are investigating the reactivity of metal-oxide nanoparticles toward Paraoxon. Previous reports have shown that magnesium oxide (MgO) and titanium oxide (TiO<sub>2</sub>) nanoparticles are effective materials for the hydrolysis and degradation of chemical warfare agents.<sup>13-16</sup> Our initial experiments have focused on determining the reactivity of MgO and TiO<sub>2</sub> nanoparticles toward Paraoxon. The hydrolysis of Paraoxon by aqueous suspensions of nanoparticles was monitored using UV-visible spectrophotometry. Magnesium oxide nanoparticles are extremely effective at degrading Paraoxon. TiO<sub>2</sub> nanoparticles are not as reactive toward Paraoxon; however, it has been reported that TiO<sub>2</sub> is an effective photocatalyst, capable of degrading chemical warfare agents. We plan to investigate this property. Using a layer-by-layer assembly method, we have attempted to immobilize



MgO and TiO<sub>2</sub> nanoparticles in polyelectrolyte thin films. Immobilized MgO nanoparticles in polyelectrolyte thin films of polyethyleneimine and polyacrylic acid have been made and characterized by atomic force microscopy. As shown in Figure 3, the surfaces are extremely rough; much rougher than would be expected for a thin film containing nanoparticles. Further investigation revealed that the commercially obtained “nanoparticles” were actually “nanostructured” microparticles. Although highly reactive due to the nanoscale features, the particles are not appropriate for making thin films according to our procedures. A new supplier has been identified and nanoparticles obtained for additional studies

## References

- (1) Yang, Y.-C.; Baker, J. A.; Ward, J. R. *Chem. Rev.* **1992**, 92, 1729.
- (2) Main, R. A.; Iverson, F. *Biochem. J.* **1966**, 100, 525.
- (3) Morales-Rojas, H.; Moss, R. A. *Chem. Rev.* **2002**, 102, 2497, and references therein.
- (4) Menger, F. M.; Tsuno, T. *J. Am. Chem. Soc.* **1989**, 111, 4903-4907.
- (5) Morrow, J. R.; Troglor, W. C. *Inorg. Chem.* **1989**, 28, 2330-2333.
- (6) Gellman, S. H.; Petter, R.; Breslow, R. *J. Am. Chem. Soc.* **1986**, 108, 2388-2394.

- (7) Kenley, R. A.; Fleming, R. H.; Laine, R. M.; Tse, D. S.; Winterle, J. S. *Inorg. Chem.* **1984**, *23*, 1870-1876.
- (8) Kazankov, G. M.; Sergeeva, V. S.; Efremenko, E. N.; Alexandrova, L.; Varfolomeev, S. D.; Rayabov, A. D. *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 3117-3119.
- (9) Kaszankov, G. M.; Sergeeva, V. S.; Borisenko, A. A.; Zatsmann, A. I.; Ryabov, A. D. *Russian Chemical Bulletin* **2001**, *50*, 1844-1848.
- (10) Berg, T.; Simeonov, A.; Janda, K. D. *J. Comb. Chem.* **1999**, *1*, 96-100.
- (11) Williams, N. H.; Takasaki, B.; Wall, M.; Chin, J. **1999**, 485-493.
- (12) Menger, F. M.; Gan, L. H.; Johnson, E.; Durst, D. H. *J. Am. Chem. Soc.* **1987**, *109*, 2800-2803.
- (13) Wagner, G. W.; Koper, O. B.; Lucas, E.; Decker, S.; Klabunde, K. J. *J. Phys. Chem. B.* **2000**, *104*, 5118-5123.
- (14) Wagner, G. W.; Bartram, P. W.; Koper, O.; Klabunde, K. J. *J. Phys. Chem. B.* **1999**, *103*, 3225-3228.
- (15) Rajagopalan, S.; Koper, O.; Decker, S.; Klabunde, K. J. *Chem. Eur. J.* **2002**, *8*.
- (16) Carnes, C. L.; Kapoor, P. N.; Klabunde, K. J. *Chem. Mater.* **2002**, *14*, 2922-2929.
- (17) Ash, B. J.; Siegel, R. W.; Schadler, L. S. *Macromolecules* **2004**, *37*, 1358-1369.
- (18) Zhang, Z.; Yang, J.-L.; Friedrich, K. *Polymer* **2004**, *45*, 3481-3485.
- (19) Lee, Y.-M.; Viswanath, D. S. *Polymer Eng. and Sci.* **2000**, *40*, 2332-2341.
- (20) Maity, A.; Biswas, M. *J. Applied Polymer Sci.* **2003**, *88*, 2233-2237.
- (21) Moustafa, A. B.; Faizalla, A.; Abd El Hady, B. M. *J. Applied Polymer Sci.* **1998**, *67*, 637-641.
- (22) Taylor, R. A.; Santora, B. P.; Gagne, M. R. "A polymer-supported rhodium catalyst that functions in polar protic solvents" *Organic Letters* **2000**, *2*, 1781-1783.
- (23) Kuchеров, A.; Finashina, E.; Kramareva, N.; Rogacheva, V.; Zevin, A.; Said-Galiyev, E.; Kustov, L. "Comparative study of Cu(II) catalytic sites immobilized onto different polymeric supports" *Macromolecular Symposia* **2003**, *204*, 175-189.
- (24) Shiraishi, Y.; Toshima, N. "Oxidation of ethylene catalyzed by colloidal dispersions of poly(sodium acrylate)-protected silver nanoclusters" *Colloids and Surfaces* **2000**, *169*, 59-66.
- (25) Vu, Y. T.; Mark, J. E. "Polymer-protected palladium nanoparticles. Morphologies and catalytic selectivities" *Colloid and Polymer Science* **2004**, *282*, 613-619.
- (26) Toshima, N.; Shiraishi, Y.; Teranishi, T. "Effect of additional metal ions on catalyses of polymer-stabilized metal nanoclusters" *Journal of Molecular Catalysis a-Chemical* **2001**, *177*, 139-147.
- (27) Cohen, R. E.; Lee, D.; Rubner, M. F. "All-Nanoparticle Thin-Film Coatings" *Nano Letters* **2006**, *6*, 2305-2312.

**VCSEL-based optical transmitter system  
using injection-locking technique**

Project Investigator:

Xiaomin Jin  
Electrical Engineering  
California Polytechnic State University  
San Luis Obispo, California

## **I. Project Overview**

This project is to study the characteristics of different semiconductor lasers with and without external injection. These lasers include Vertical-Cavity Surface-Emitting (VCSEL) Lasers, Distributed Feedback (DFB) Laser, and Fabry Perot (FP) Laser. Each type of the laser has one or more of the following measurements controlled and recorded via a PC with LabVIEW 7.1 Student Edition: DC Light-Voltage-Current (LVI) curves, optical spectrums, bandwidths, and also relative intensity noise (RIN).

Experimental work at Cal Poly are: The DC LVI curves show the operating characteristics of the laser diodes with the threshold currents that indicating simulated emission at different temperatures. The LVI curves are two sets of plots with a given set of input current source versus the laser optical power outputs and the voltages of the laser diodes. The current versus voltage measurement is only available for the VCSELs. The optical power spectrums of the lasers are measured with HP71450A Optical Spectrum Analyzer. The ILX-Lightwave LDC-3744B Laser Diode Controller acts as a current source and a temperature controller together with the ILX Temperature Controlled Laser Diode Mount for the VCSELs measurement. Then there is the bandwidth measurement which uses the HP8720B Network Analyzer, frequencies range from 130 MHz to 20 GHz to send out a pulse to modulate the VCSELs or the FP lasers. A high speed photo-detector is to detect the input modulated laser signal and then amplified by a broadband amplifier before going to port 2 of the network analyzer. Finally, the relative intensity noise (RIN) is to measure how much noise (dB) per frequency (Hz) for the lasers.

During this year, PI work with other university and research institutes and published several papers on laser injection and integrate optoelectronic device design. PI also submitted several proposes externally related with this or use this research results.

## **II. Publication under support of the project:**

- X. Jin and S. L Chuang, "Bandwidth Enhancement of Fabry-Perot Quantum-well Lasers by Injection-locking", *Solid State Electronics*, vol. 50, Issue 6, pp. 1141-1149, June 2006.
- X. Jin and S. L. Chuang "Injection-Locking in Fabry-Perot Quantum-well Lasers", 2005 International Semiconductor Device Research Symposium (ISDRS 2005), Bethesda, MD. Dec 6-9<sup>th</sup> 2005.
- Jin, A. Hsu, and S. L. Chuang, "Study of Optical-feedback using an Integrated Laser-modulator/amplifier Device", *2006 Integrated photonics research and application topical meeting* (IPRA 2006), Uncasville, CT, USA , April 24-26<sup>th</sup>, 2006

## **III. Proposal submitted related to the result of the project:**

- Proposal for NSF05-759 CAREER: “**High-speed all-optical wavelength conversion using injection-locking**”, X. Jin, California Polytechnic State University, Department of Electrical Engineer, San Luis Obispo, California, submitted 7/16/06, \$420,426. (pending) (This is a direct continuous of the current project)
- ASEE: The Air Force Summer Faculty Fellowship Program (SFFP) 2007, “**Enhance dynamic property of semiconductor laser through injection-locking technology**”, X. Jin, California Polytechnic State University, Department of Electrical Engineer, San Luis Obispo, California, submitted white paper (15pages) 9/28/06 to Dr. Thomas R. Nelson, Jr. and Dr. Vassilios Kovanis, AFRL, SENSORS DIRECTORATE, Electro-Optics Components Branch, 2241 Avionics Circle, Wright Patterson AFB, Ohio 45433-7320. They are very interested in the ideal and our research outcome. We work together and refine the injection-locking ideal. In Nov, 2006 submitted the 4 page proposal to ASEE web-site (pending). (This is a direct continuous of the current project and external collaboration)
- Proposal for Agilent Research Grant 2006 “**Modeling and testing of semiconductor lasers, cables, and photodiodes for interferometer measurement system applications**”, X. Jin, X. Jin, California Polytechnic State University, Department of Electrical Engineer, San Luis Obispo, California, Submitted Feb. 2006, \$78,100. (funded). (This is an indirect outcome of the current project, which use the several setups we developed in this project)
- Proposal for JDSU of industry-university collaboration and donation “**Multimode fiber bandwidth measurement**” and “Donation of transceivers and high-speed laser”, X. Jin, X. Jin, California Polytechnic State University, Department of Electrical Engineer, San Luis Obispo, California, Submitted 11/30/2006. We talk about this collaboration over the phone several times, it is very promising. (This is an indirect outcome of the current project, which use the setup we developed in this project Fig IV-14)

#### IV. Technical results of the project:

##### 1. Different Types of Lasers

###### A. Vertical-Cavity Surface-Emitting Lasers (VCSELs)

We tested three types of VCSELs with wavelength 850nm to measure the characteristics. The three types of VCSELs are low-power, default, and high-power.

###### B. Distributed Feedback (DFB) Laser

This is a 1550nm laser. Because this laser has its own current and temperature controller, and the temperature controller does not work; therefore, the LVI curves respect to different temperature are not measured.

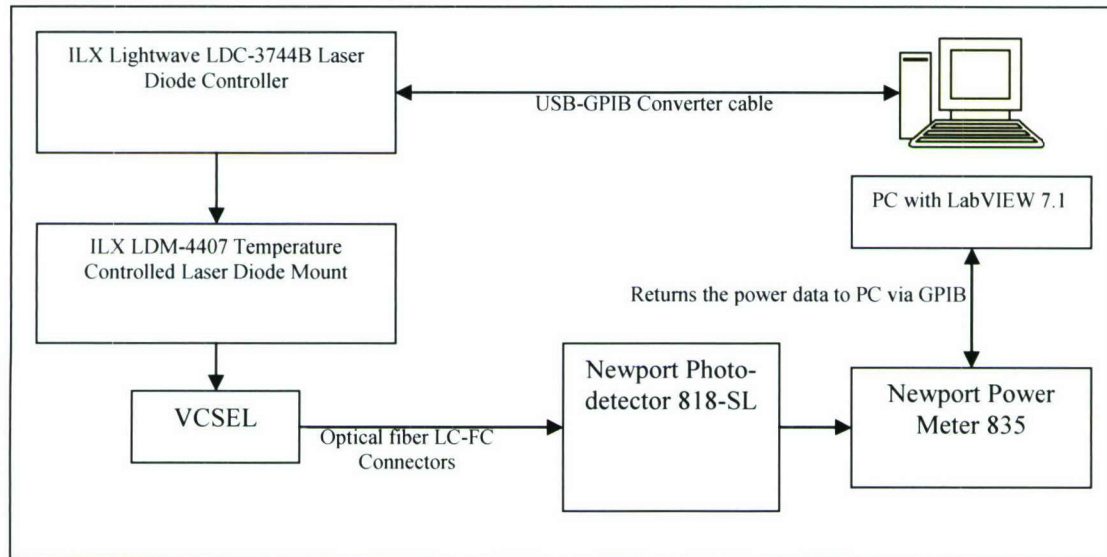
### C. Fabry Parot Laser

This type of laser is powered by an external amplifier with internal current source.

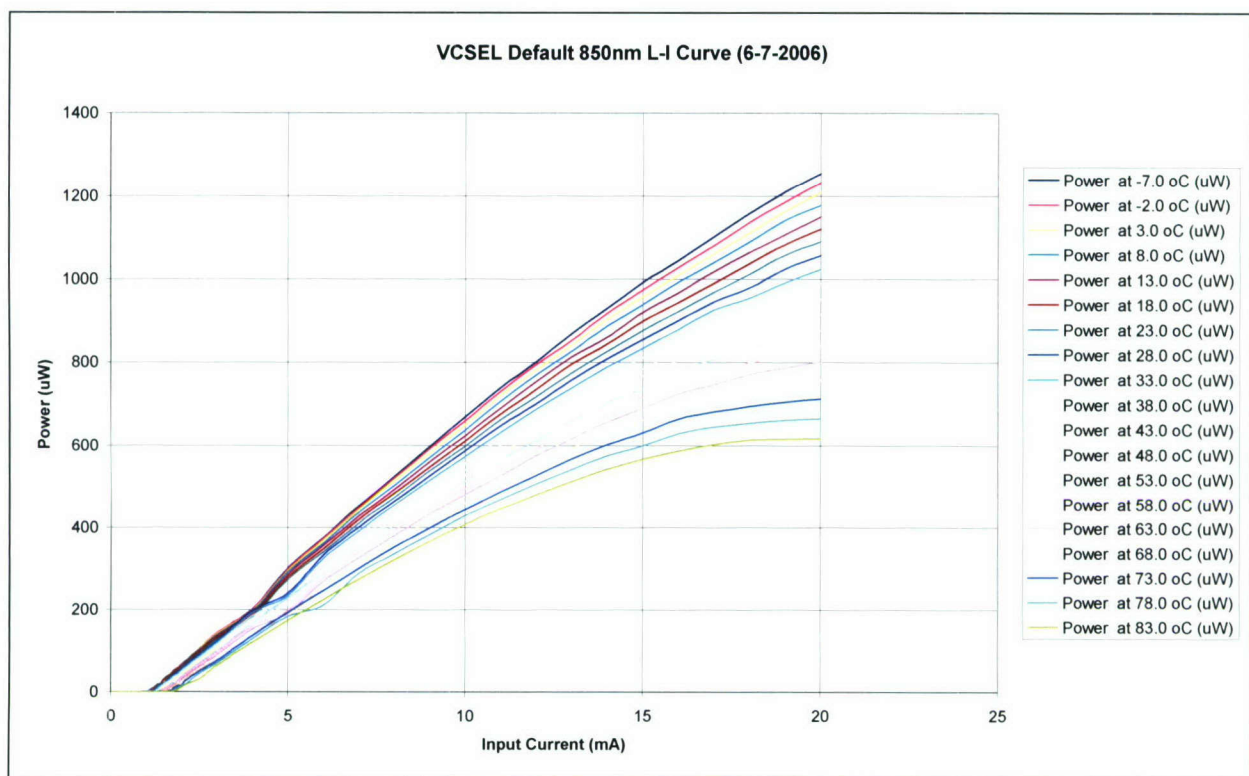
## 2. Experiments:

### A. DC LVI Curves

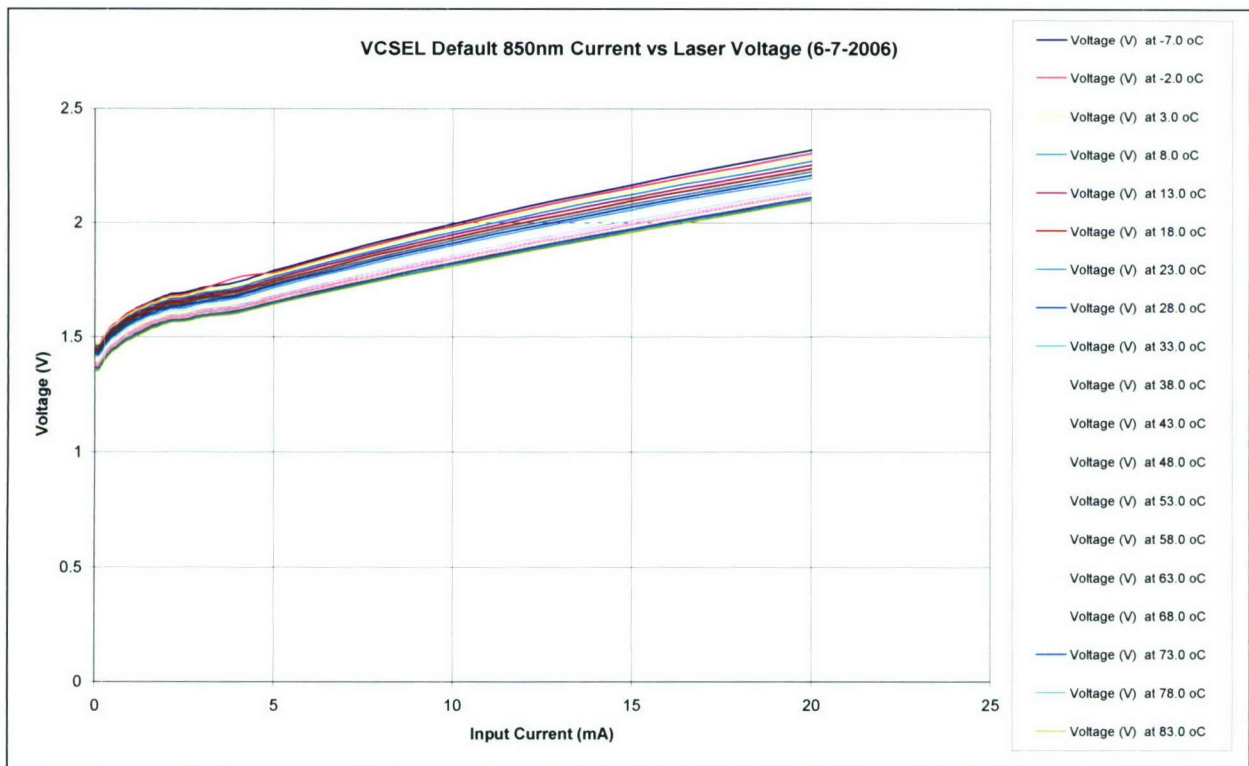
#### 1) VCSEL



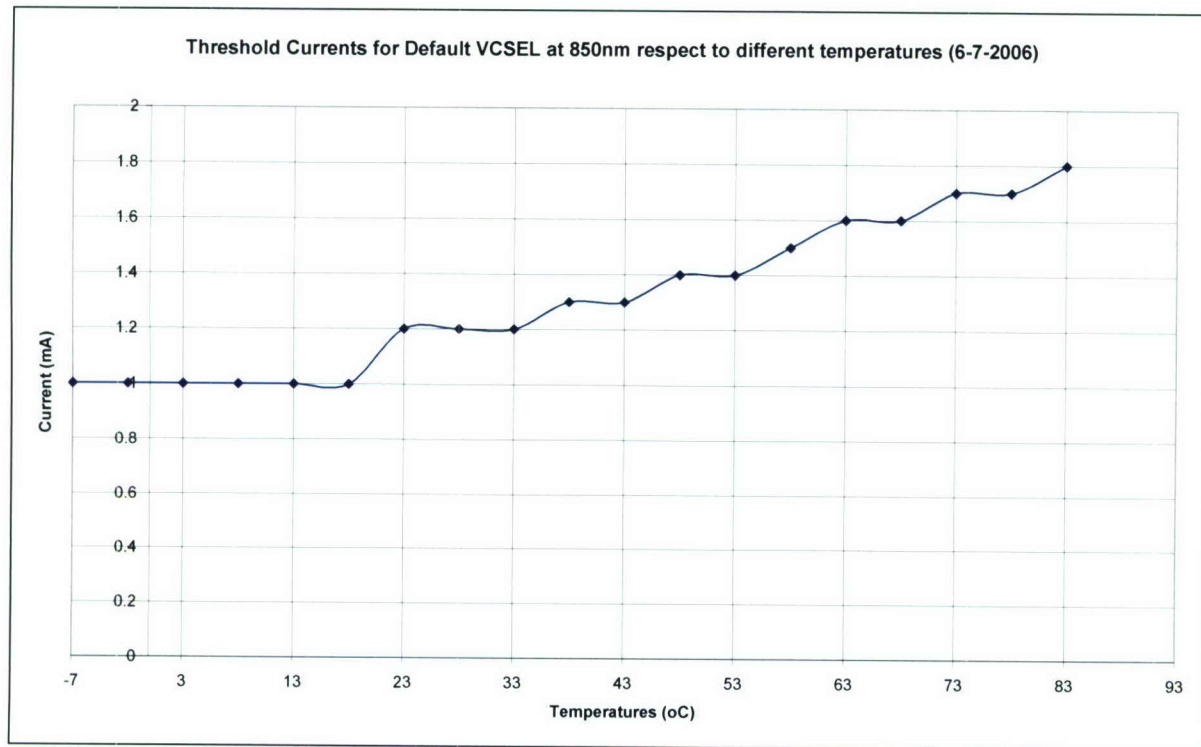
**Figure IV-1:** Block diagram for VCSEL DC LVI Curves



**Figure IV-2: LI Curve for Default VCSEL from -7 to 83 °C**



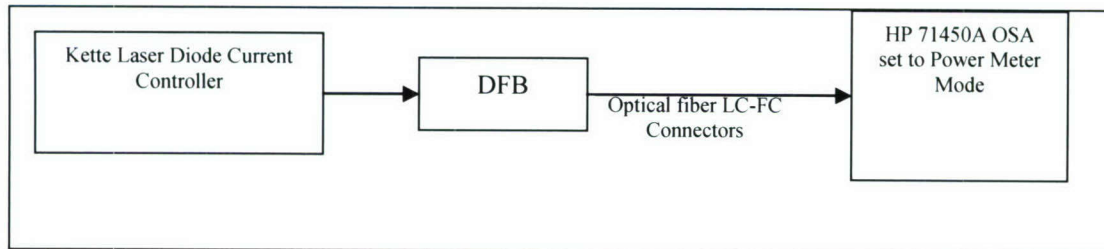
**Figure 2-3: VI Curve for Default VCSEL from -7 to 83 °C**



**Figure IV-4:** *Threshold Currents Plot for Default VCSEL from -7 to 83 °C*

## 2) DFB Laser

This DFB laser has its own set of current source and the temperature controller does not work. Moreover, there are also no external connectors for connecting a digital multi-meter to measure the voltages of the laser; therefore, only the LI curve respect to the room temperature can be recorded.



IV-5: Block diagram for DFB Laser LI curves

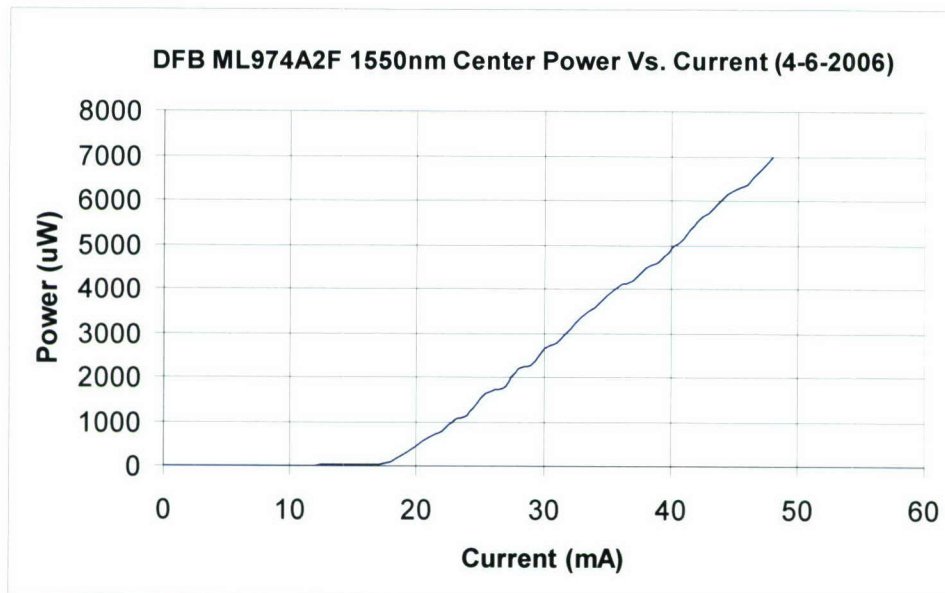


Figure IV-6: LI Curve for DFB Laser

## B. Optical Power Spectrums

### 1) VCSEL

The optical power spectrums for VCSELs changes with different temperatures at different current inputs. The LabVIEW program file name “**LaserCtrl-OSA-NTA-RFSP .vi**” automatically controls the ILX LDM-3744B temperature - current controller and the HP 71450A optical spectrum analyzer (OSA) to record different power spectrums at a given range of input currents and temperatures. The program then finally outputs the entire data base of optical power spectrums at different conditions then generates two separate spreadsheets: 1) constant temperature with varying input currents and 2) constant input current with varying temperatures. User can later plot out different power spectrum combinations by using

“Reading indexed data v1.0.vi” LabVIEW program together with the data base files and provide the parameters for the constant temperature and current input.

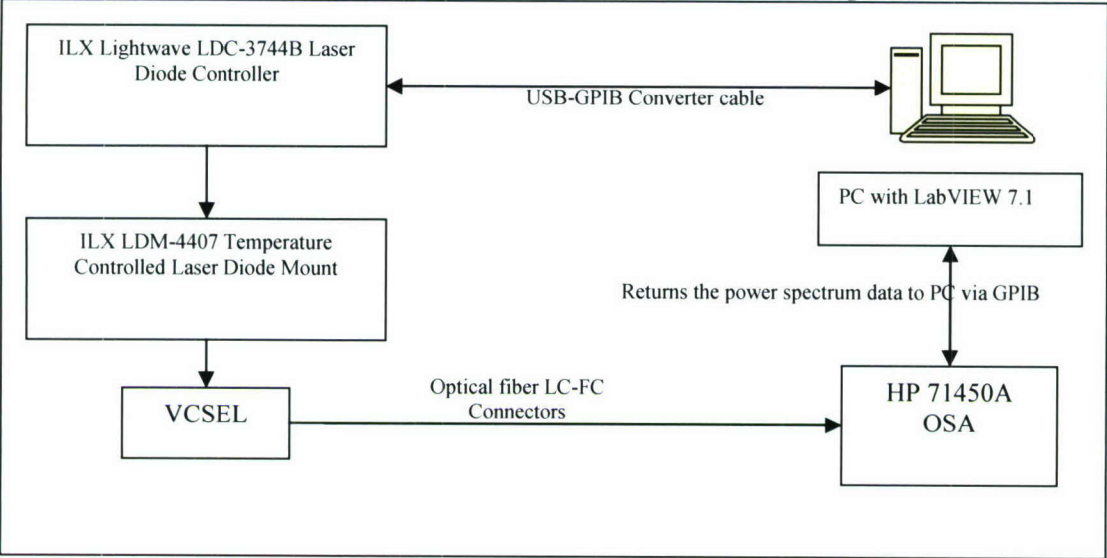


Figure IV-7: Block Diagram for VCSEL Optical Power Spectrum Measurement

Before recording the power spectrum, the HP 71450A OSA has to be set with the following parameters:

Resolution Bandwidth (nm):	0.08
Span (nm):	10
Sensitivity (dBm):	-75

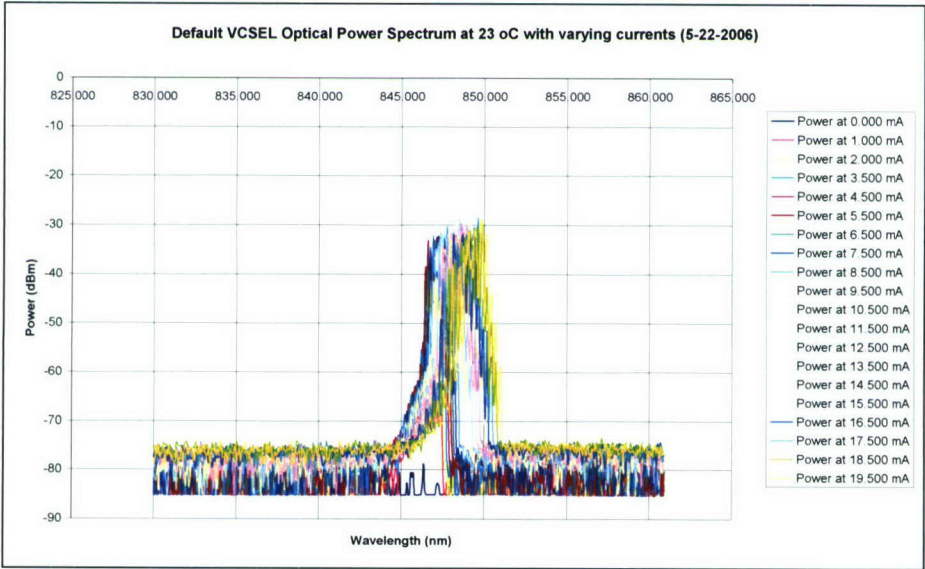


Figure IV-8: Optical Power Spectrums at varying input currents for Default VCSEL with constant 23°C. The center wavelength shifts to the right as the input current increases.

## 2)DFB Laser

Because this laser’s temperature controller does not work; therefore, there is only optical

power spectrum at room temperature around 23°C

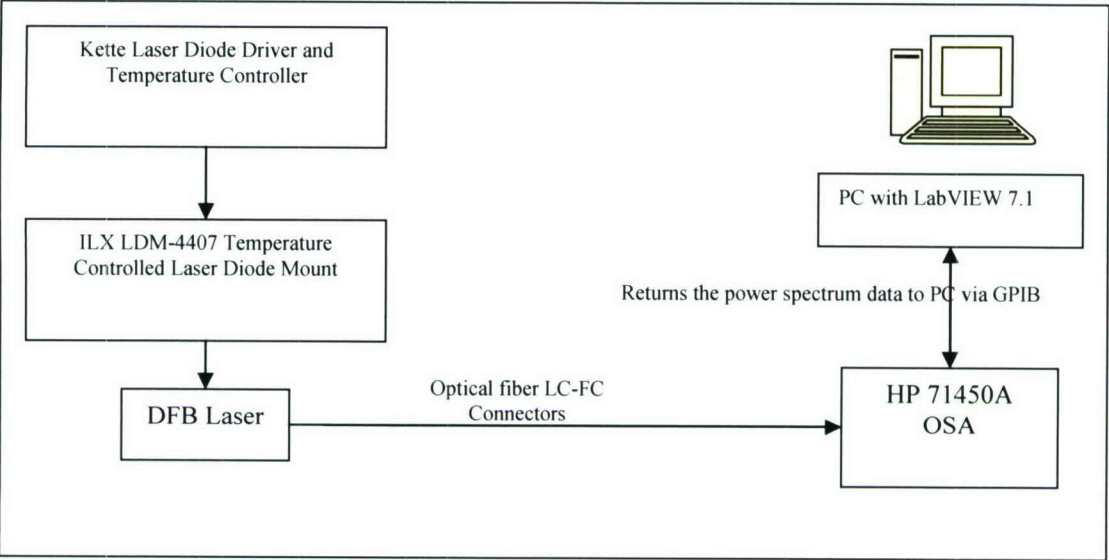


Figure IV-9: Block diagram for DFB Laser optical power spectrum measurement.

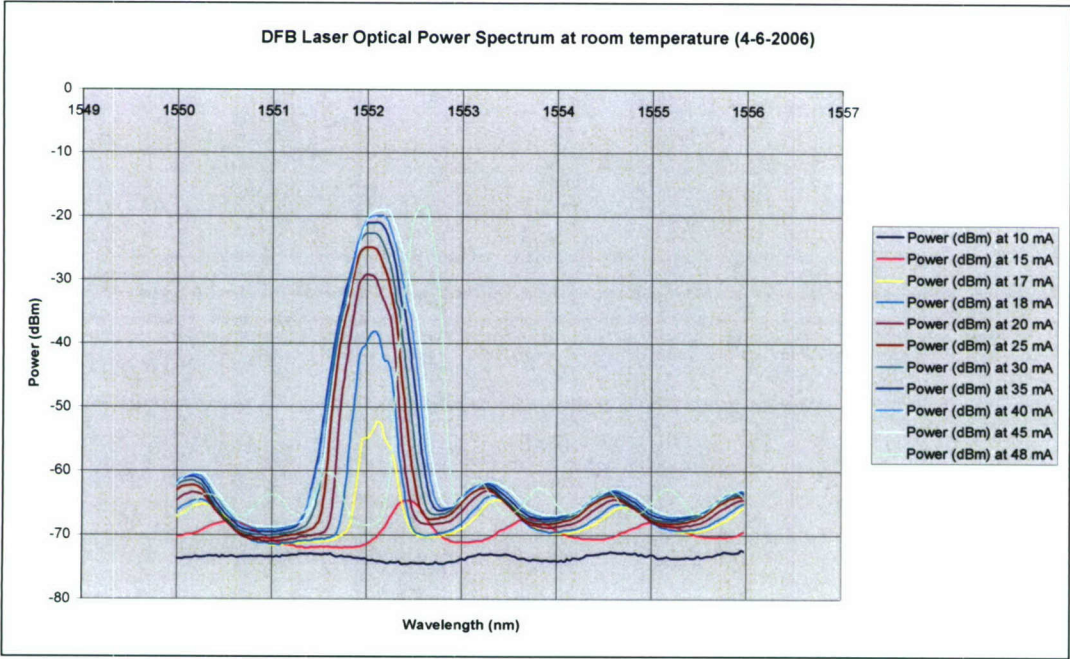
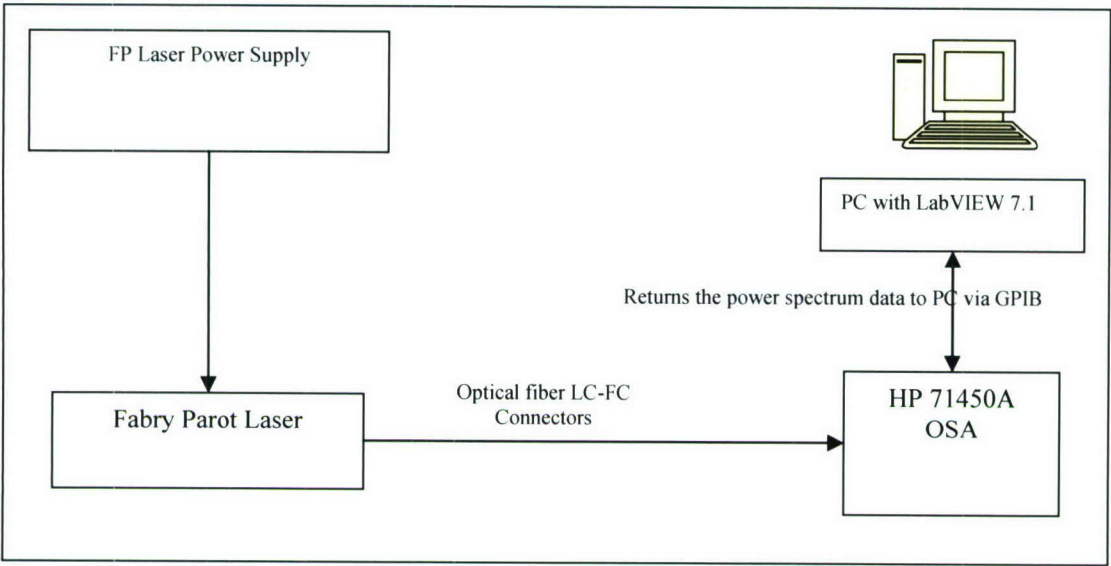


Figure IV-10: Optical Power Spectrum for Distributed Feedback Laser

The optical power spectrum again indicates as the input current increases, the center wavelength, which is the peak of each plot in the graph above, shifts to the right.

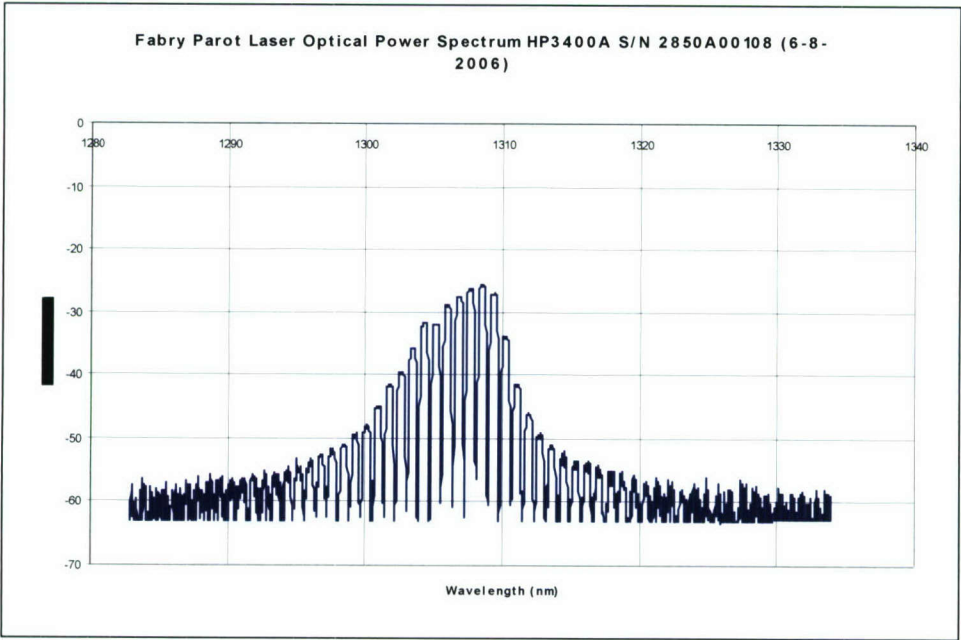
### 3) Fabry Parot Lasers



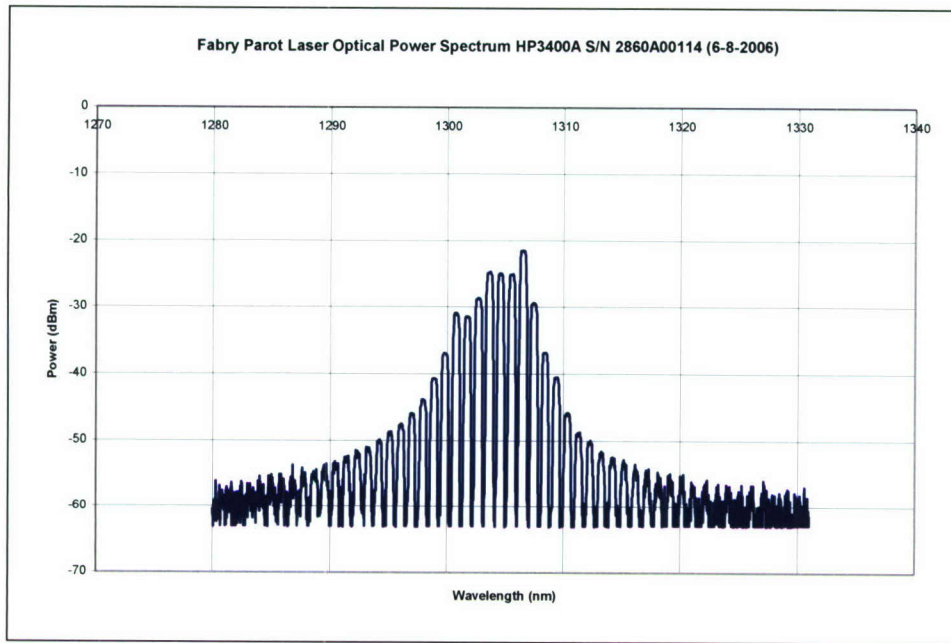
**Figure IV-11: Block diagram for Fabry Parot Laser to measure the optical power spectrum.**

Before recording the power spectrum, the HP 71450A OSA has to be set with the following parameters:

	#1	#2	#3	#4
<b>Resolution Bandwidth (nm):</b>	0.5	0.5	0.5	0.5
<b>Span (nm):</b>	50	50	50	40
<b>Sensitivity (dBm):</b>	2.82	2.82	2.82	2.74



**Figure IV-12: Optical Power Spectrum for FP Laser #1**

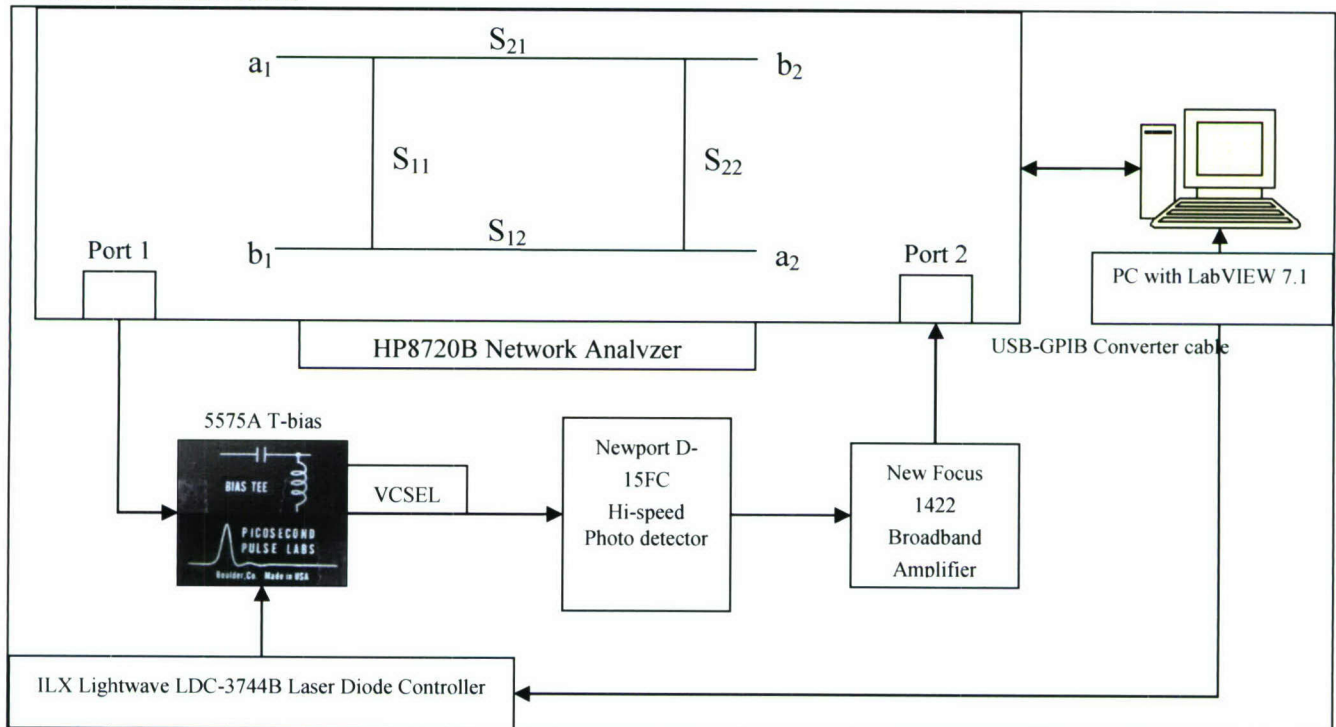


**Figure IV-13:** Optical Power Spectrum for FP Laser #2

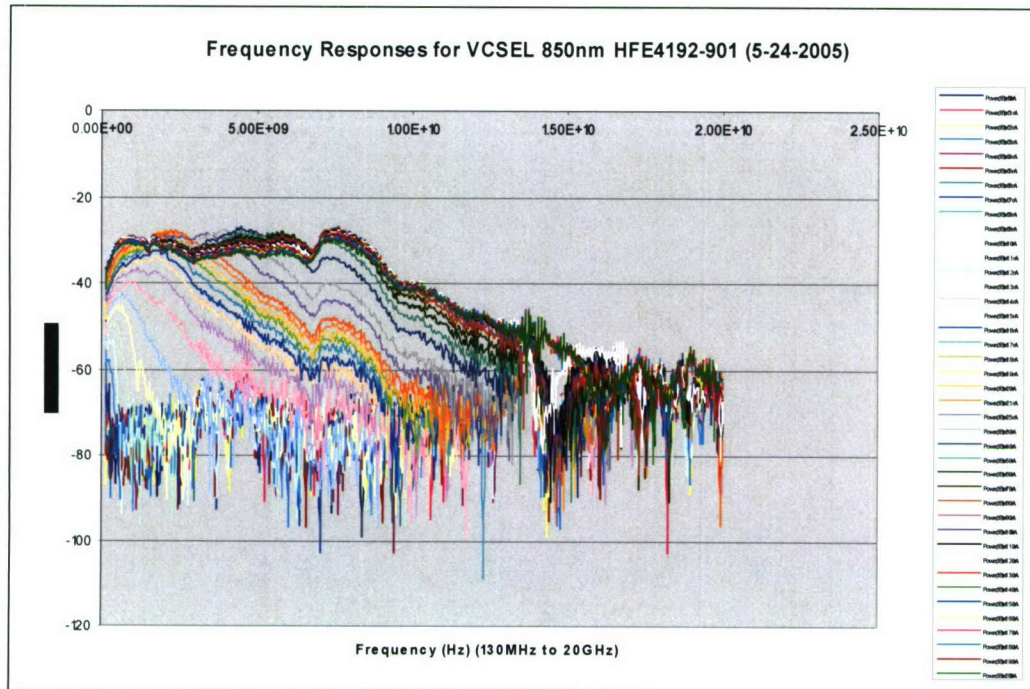
### C. Bandwidths

For bandwidth measurement, there are only VCSELs and Fabry Parot lasers, because DFB laser cannot be modulated with the network analyzer.

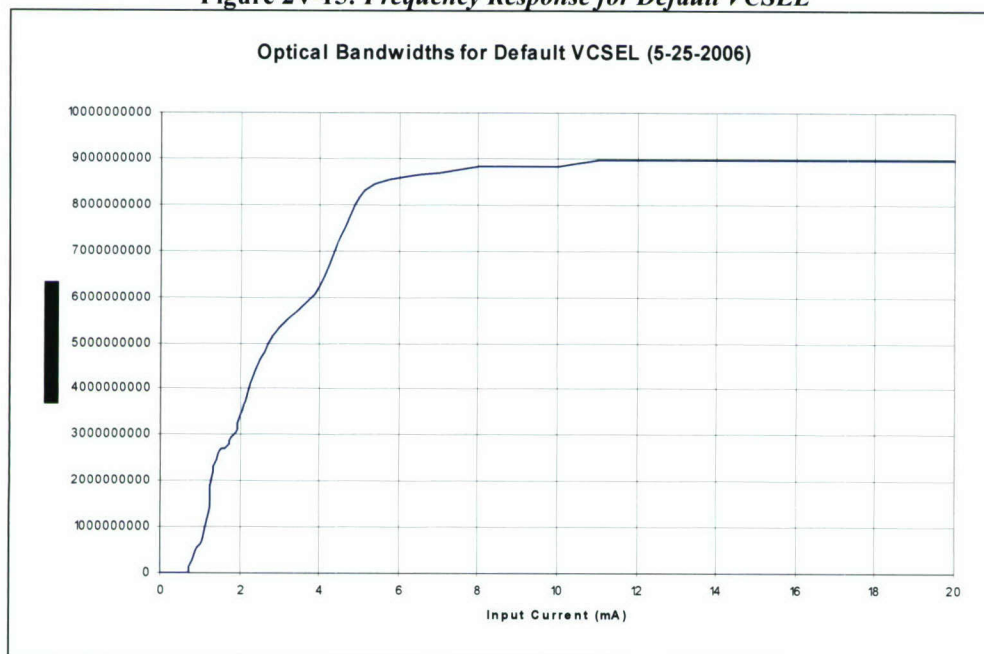
#### VCSELs Bandwidths



**Figure IV-14:** Block Diagram Showing how to obtain the electrical bandwidth of the VCSELs from the network analyzer.



**Figure 2V-15: Frequency Response for Default VCSEL**



**Figure IV-16: Bandwidths from different input current (0 mA to 20 mA). The Default VCSEL can go up to 9 GHz.**

#### D. Relative Intensity Noise (RIN)

It is to measure how noisy the laser is in dB/Hz ( $10 \log (RIN / \text{unit BW})$ ). However, the HP 8593A Spectrum Analyzer cannot return the data in unit *dB/Hz*; therefore, a special experimental setup and post processing are required to obtain the final RIN values.

#### FP and DFB Laser RIN

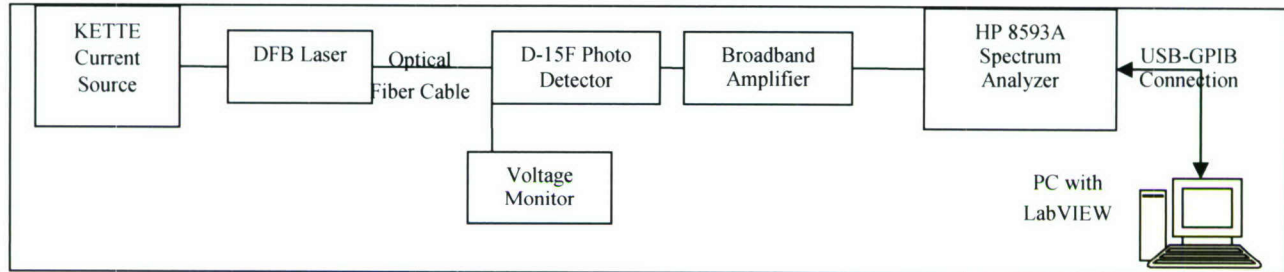


Figure IV-17: Block Diagram for measuring DFB laser relative intensity noise (RIN)

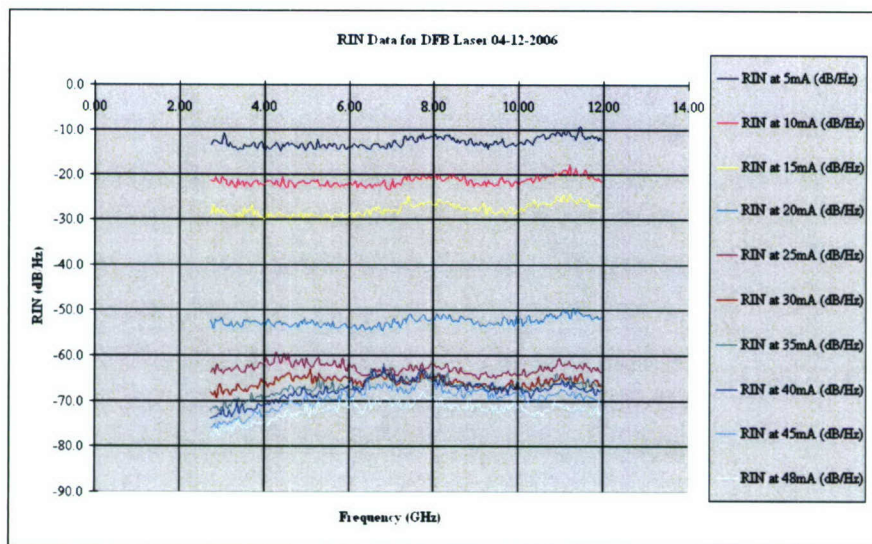


Figure IV-18: Calculated Relative Intensity Noise for DFB laser

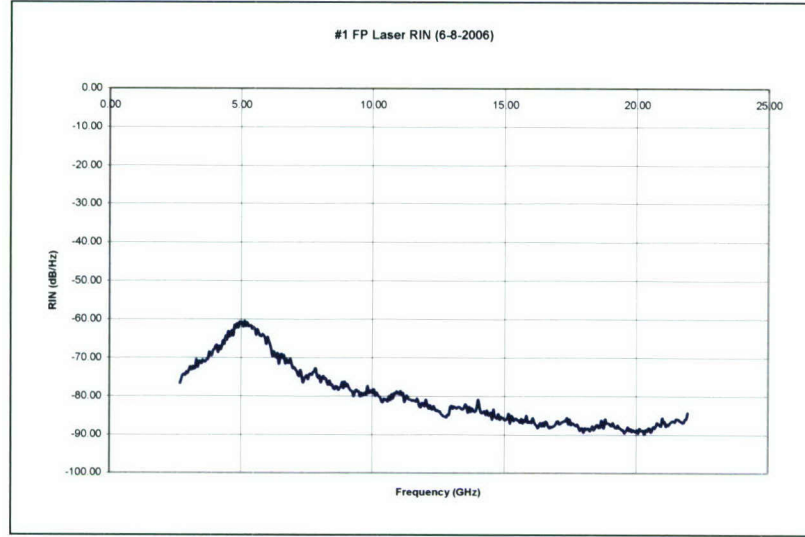


Figure IV-19: RIN data for #1 FP Laser

### 3. Theoretical work: Improved Injection-locking theory for multi-mode lasers, such as FPs or VCSELs

We consider two semiconductor lasers with a small difference in resonant wavelength (or frequency). The pump light is injected from one laser, the single-mode master laser, into the other laser, referred to as the test laser or slave laser. The reverse light path is blocked by an optical isolator. The locking occurs within a certain frequency locking range defined by the injection level and the amplitude-phase coupling coefficient. It is assumed that the effects of spatial hole-burning in a FP laser are negligible. The theoretical analysis describing the injection-locked test laser is based on the rate equations, shown as below. We also need to consider the possible excitation of the neighboring longitudinal mode with different injections. This can be accounted for by a rate equation for the photon density  $S_u$  in the side modes or the group of side modes, which represents the unlocked photon density. Therefore, the total photon density is  $S = S_u + S_l$ , where  $S_l$  is the photon density in the locked mode.

$$\begin{aligned} \frac{dS_l(t)}{dt} &= \Gamma_l v_g S_l(t) G_l [1 - \varepsilon_l S_l(t) - \varepsilon_u S_u(t) - \varepsilon_i S_i] \\ &- \frac{S_l}{\tau_{pl}} + 2k_c \sqrt{S_l(t) S_i} \cos(\phi_l(t) - \phi_i) \end{aligned} \quad (1)$$

$$\begin{aligned} \frac{dS_u(t)}{dt} &= \Gamma_u v_g S_u(t) G_u [1 - \varepsilon_l S_l(t) - \varepsilon_u S_u(t) - \varepsilon_i S_i] \\ &- \frac{S_u(t)}{\tau_{pu}} \end{aligned} \quad (2)$$

$$\begin{aligned} \frac{d\phi_l(t)}{dt} &= -k_c \sqrt{\frac{S_l}{S_i}} \sin(\phi_l(t) - \phi_i) - (\omega_l - \omega_0) \\ &+ \frac{\alpha}{2} \Gamma_l v_g g'_l [N(t) - N_0] [1 - \varepsilon_l S_l(t) - \varepsilon_u S_u(t) - \varepsilon_i S_i] \end{aligned} \quad (3)$$

$$\frac{dN(t)}{dt} = \frac{I(t)}{eV} - \frac{N(t)}{\tau_n}$$

$$\begin{aligned}
& -v_g S_l(t) G_l [1 - \varepsilon_l S_l(t) - \varepsilon_u S_u(t) - \varepsilon_i S_i] \\
& -v_g S_u(t) G_u [1 - \varepsilon_l S_l(t) - \varepsilon_u S_u(t) - \varepsilon_i S_i]
\end{aligned} \quad (4)$$

where  $S_i$  is injected photon density into the test laser cavity from the master laser,  $\phi_l(t)$  and  $\phi_i(t)$  are the phase of the locked mode and injected laser field,  $N(t)$  is the carrier density,  $\omega_l$  is the master laser frequency,  $\omega_0$  is the cavity resonance frequency of the test laser,  $\alpha$  is the linewidth enhancement factor,  $k_c = \frac{c}{2n_g L}$  is the coupling coefficient,  $n_g$  is the group index of the

test laser,  $L$  is the test laser cavity length, the  $I(t)$  is the test laser current,  $V$  is the volume of the active region,  $e$  is the unit charge of the carrier,  $\tau_n$  is the carrier lifetime,  $v_g$  is the group velocity,  $\tau_{pl}$  and  $\tau_{pu}$  are the photon lifetime for locked mode and unlocked mode, and  $G_l = g_{l0} + g'_l (N(t) - N_0)$  and  $G_u = g_{u0} + g'_u (N(t) - N_0)$  are the gain of the locked and unlocked laser, and  $g'_l$  and  $g'_u$  are the differential gain. For the small-signal analysis, the quantity  $N(t) - N_0$  will equal the small-signal change in carrier density, denoted by  $n$ . The spontaneous emission term has been neglected because the test laser is biased above threshold. We also include both nonlinear gain saturation coefficients,  $\varepsilon_l$  and  $\varepsilon_u$ , as an improvement of a previous model for injection locking in semiconductor QW lasers. Our contributions are listed below. First, we include, the optical confinement factor  $\Gamma$ , which is an important parameter for SCH QW laser structures, in the rate equations. The factor is well-known to be important for separate confinement quantum-well structures, however, it is usually ignored in the literature on injection locking. Second, we include the nonlinear gain saturation of the slave laser due to the injected photon density  $\varepsilon_i$  and all optical modes in the cavity, which has been used in earlier studies on high-speed lasers where the gain of the test laser light is suppressed due to the presence for the injected light. This is an important phenomenon if we inject light in the laser gain region and it cannot be neglected. As we will show later, inclusion of  $\Gamma$  and  $\varepsilon_i$  is important to extra consistent parameters for the gain and differential gain for quantum-well lasers.

If no light is injected into the test laser (free-running  $S_i=0$ ), we can see from Eq. (1)-(4) that only the photon density and carrier density are coupled. The phase term is not necessary for solving the photon density and carrier density. However in an injection-locked laser, whose injected photon density  $S_i$  is non-zero, there is a coupling of the magnitude and phase of photon field and carrier density through externally injected light from the master laser. This is a unique and important characteristic of injection-locked semiconductor lasers called amplitude and phase coupling.

By solving for the steady-state solution of the rate equations, we obtain the expression of the locking bandwidth which is given as a function of the injection rate and the phase difference

$$\begin{aligned}
\Delta\phi &= \phi - \phi_i \\
\Delta\omega &= \omega_l - \omega_0 = \frac{c}{2n_g L} \sqrt{\frac{S_l}{S}} \left[ \sin(\Delta\phi) - \frac{\alpha}{2} \cos(\Delta\phi) \right]
\end{aligned} \quad (5)$$

There are two ways by which a laser field oscillating at cavity resonance frequency  $\omega_0$  in the absence of injection can be forced instead to oscillate at the frequency of the master laser  $\omega_l$ . The first mechanism which appears in every locking system is to have the injected field add

an out-of-phase component to the lasing mode of the test laser. This alters the phase of the test laser field and changes the frequency of the test laser field. The first term on the right-hand side of Eq. (5) represents this frequency shift. The second way to have the test laser field oscillate at  $\omega_i$  is to alter the cavity resonance frequency  $\omega_0$  by the injected field changing the gain required to maintain the steady-state test laser intensity. The change in gain is accompanied by a shift in cavity resonance frequency. This is an extra term peculiar to semiconductor lasers arising from the refractive index dependence on carrier density. The second term of the expression (5) accounts for this cavity frequency shift due to the refractive index dependence on gain change mechanism, characterized by the linewidth enhancement factor  $\alpha$ .

The injection-locking range is determined by both detuning ( $\Delta\omega = \omega_i - \omega_0$ ) and the injection power. Using Eq. (5) and the condition that the injection locking is a constructive interaction with regard to the power balance, we obtain an asymmetric locking bandwidth [22]

$$\Delta\omega_{\min} = -\frac{c}{2n_g L} \sqrt{\frac{S_i}{S}} (1 + \alpha^2) < \Delta\omega < -\frac{c}{2n_g L} \sqrt{\frac{S_i}{S}} = \Delta\omega_{\max} \quad (6)$$

Outside this region is the unlocked region, where the injection level is too low or the detuning is too high to reach locking condition. A detailed stability analysis can be found in, which also calculates a self pulsation zone (a Hopf bifurcation), chaotic zone, and a coherence collapse zone. Here, we focus on the stable locking zone. We carefully control our experiment to make sure that the injected power is not strong enough to cause self pulsation.

We also derive the small-signal modulation response of an injection-locked laser from rate equations, Eq. (1)-(4). The changes in the lasing mode photon densities and carrier density due to small-signal modulation are assumed to be much smaller than the steady-state value of the photon and carrier densities. To solve for the small-signal modulation response, the expressions for carrier and photon densities are

$$N(t) = N_0 + n(\omega)e^{j\omega t} \quad (7)$$

$$S_l(t) = S_{l0} + s_l(\omega)e^{j\omega t} \quad (8)$$

$$S_u(t) = S_{u0} + s_u(\omega)e^{j\omega t} \quad (9)$$

$$\phi_l(t) = \phi_{l0} + \phi_l(\omega)e^{j\omega t} \quad (10)$$

$$I(t) = I_0 + i(\omega)e^{j\omega t} \quad (11)$$

In our notation, capital letters stand for steady-state values. Taylor's series expansion is used to simplify the small-signal form of the rate equations. Terms containing products of steady-state and small-signal components are linearized, and only the first-order terms are retained. The small-signal rate equations can be expressed as follows:

$$\begin{bmatrix} -j\omega & & & \\ & -j\omega & & \\ & & -j\omega & \\ & & & -j\omega \end{bmatrix} \begin{bmatrix} n(\omega) \\ s_l(\omega) \\ s_u(\omega) \\ i(\omega) \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ -\frac{i(\omega)}{eV} \\ 0 \end{bmatrix} \quad (12)$$

where

$$\begin{aligned} a_{11} &= -\Gamma_l v_g g_{l0} (2\varepsilon_l S_{l0} + \varepsilon_u S_{u0} + \varepsilon_i S_i - 1) - \frac{1}{\tau_{pl}} + k_c \sqrt{\frac{S_i}{S_{l0}}} \cos(\phi_{l0} - \phi_i) \\ a_{12} &= -\Gamma_l v_g g_{l0} \varepsilon_u S_{l0} \\ a_{13} &= \Gamma_l v_g g'_{l0} S_{l0} (1 - \varepsilon_l S_{l0} - \varepsilon_u S_{u0} - \varepsilon_i S_i) \\ a_{14} &= -2k_c \sqrt{S_{l0} S_i} \sin(\phi_{l0} - \phi_i) \\ a_{21} &= -\Gamma_u v_g g_{u0} \varepsilon_l S_{u0} \\ a_{22} &= -\Gamma_u v_g g_{u0} (\varepsilon_l S_{l0} + 2\varepsilon_u S_{u0} + \varepsilon_i S_i - 1) - \frac{1}{\tau_{pu}} \\ a_{23} &= \Gamma_u v_g g'_{u0} S_{u0} (1 - \varepsilon_l S_{l0} - \varepsilon_u S_{u0} - \varepsilon_i S_i) \\ a_{24} &= 0 \\ a_{31} &= v_g g_{l0} (2\varepsilon_l S_{l0} + \varepsilon_u S_{u0} + \varepsilon_i S_i - 1) + v_g g_{u0} \varepsilon_l S_{u0} \\ a_{32} &= v_g g_{u0} (\varepsilon_l S_{l0} + 2\varepsilon_u S_{u0} + \varepsilon_i S_i - 1) + v_g g_{l0} \varepsilon_u S_{l0} \\ a_{33} &= -\frac{1}{\tau_n} - v_g g'_{l0} S_{l0} (1 - \varepsilon_l S_{l0} - \varepsilon_u S_{u0} - \varepsilon_i S_i) - v_g g'_{u0} S_{u0} (1 - \varepsilon_l S_{l0} - \varepsilon_u S_{u0} - \varepsilon_i S_i) \\ a_{34} &= 0 \\ a_{41} &= \frac{k_c}{2S_{l0}} \sqrt{\frac{S_i}{S_{l0}}} \sin(\phi_{l0} - \phi_i) \\ a_{42} &= 0 \\ a_{43} &= \frac{\alpha}{2} \Gamma_l v_g g'_{l0} (1 - \varepsilon_l S_{l0} - \varepsilon_u S_{u0} - \varepsilon_i S_i) \\ a_{44} &= -k_c \sqrt{\frac{S_i}{S_{l0}}} \cos(\phi_{l0} - \phi_i) \end{aligned}$$

After eliminating the carrier density  $n$ , the phase term  $\phi_l$  and solving for  $s = s_l(\omega + s_u(\omega))$ , the intrinsic modulation response is obtained  $|M_{im}(\omega)|^2 = [S_l(\omega) + S_u(\omega)]^2 / i(\omega)^2$ . It is very important to note that the electrical modulation is an extrinsic modulation which includes parasitic and transport effects. Parasitic effects come from the bias circuit and the shunting of modulation current around the active layer, which will cause a low frequency roll-off of modulation response. At the same time, the carriers are injected from the outer edge of SCH region. The injected carriers diffuse through the SCH region and are captured into the quantum wells before recombining by stimulated emission processes. In the separate confinement structure QW lasers, the carrier diffusion, capture, and escape into and from QWs are usually defined to characterize the carrier transport processes which give a parasitic-like roll-off and are indistinguishable from parasitic effects. In quantum-well lasers, the carrier transport time is an important limit for MQW laser modulation bandwidth. To complete the theory, the final modulation response of injection-locked lasers is

$$\begin{aligned}
|M(\omega)|^2 &= \frac{1}{1 + (\omega / \omega_p)^2} |M_{\text{int}}(\omega)|^2 \\
&= \frac{1}{1 + (\omega / \omega_p)^2} \left| \frac{S_l(\omega) + S_u(\omega)}{i(\omega)} \right|^2
\end{aligned} \tag{13}$$

where  $\omega_p = 2\pi f_p$  is the low frequency roll-off due to the transport effect and parasitic effect. The calculation result is shown below.

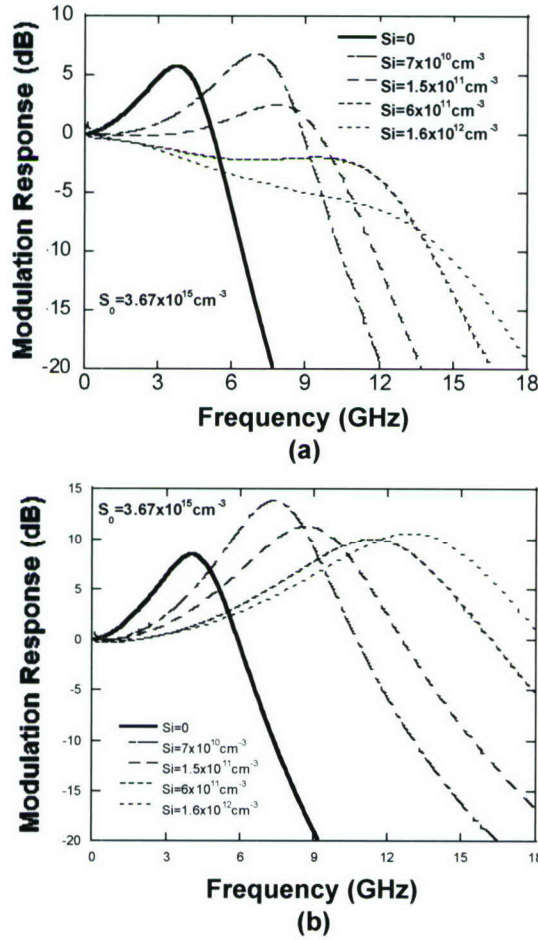


Fig. IV-20. Theoretical calculation of the small-signal modulation response of an injection-locked test laser with a fixed test laser photon density  $S_0 = 3.6 \times 10^{15} \text{ cm}^{-3}$  for 30mA test laser bias. The injected photon number varies from zero (free-running) to  $S_i = 1.2 \times 10^{12} \text{ cm}^{-3}$  and is proportional to the injection power. (a) The low frequency roll-off is included with  $f_p = 7 \text{ GHz}$ . (b) The low frequency roll-off is not included with  $f_p = \infty \text{ GHz}$ .

## **Nonlinear constitutive equations for cartilage growth analysis**

Project Investigators:

Steve Klisch and Andrew Davol  
Mechanical Engineering  
California Polytechnic State University  
San Luis Obispo, California

The articular cartilage (AC) of human joints can tolerate years of repetitive loading; however, AC damage and degeneration occur often with abnormal mechanical loading conditions, traumatic joint injury, and arthritis. AC has a poor intrinsic healing capacity that makes the repair of damaged and degenerated cartilage a formidable task. The long-term goal of the research described here is to develop and use analytical and experimental methods to quantitatively characterize the complex relationship between stimuli, cellular metabolism, tissue growth, and biomechanical function for AC.

One of the most promising applications for tissue-engineered AC constructs is in the treatment of arthritis. Arthritis is estimated to affect one in six Americans and, as the population ages, arthritis will be a major concern for the health care system in the U.S. Studies have indicated that veterans are at an increased risk of arthritis; a possible explanation for this is that one of the most common forms of arthritis is osteoarthritis which has been linked to traumatic joint injury. Also, studies have linked rheumatoid arthritis (a systemic autoimmune disease) to post traumatic stress disorder in Vietnam Veterans.

In our 2005 C3RP proposal, we stated that the results would be used as preliminary studies for future research grant proposals. The results were used as preliminary data for two grant proposals: an NSF Research in Undergraduate Institutions (RUI) proposal titled Analytical and experimental study of articular cartilage growth in vitro, and an NIH Academic Research Enhancement Award (AREA) proposal titled A continuum model for cartilage growth biomechanics. The NSF RUI proposal was funded and the NIH RUI proposal is pending.

The overall aim from our 2005 C3RP proposal was to ***develop a continuum mechanics model for the proteoglycan stress constitutive equation using a physiological model to calculate an effective fixed charge density***. Specific aims from our 2005 C3RP proposal are stated below (*in italics*), followed by the related research findings. References are made to the publications supported by this C3RP grant, listed at the end of this report.

***Specific Aim 1:*** *For all specimens, the data obtained for the masses of GAG, COL, and water will be used to calculate the apparent constituent densities,  $\rho^p, \rho^c, \rho^w$  (per tissue volume). Confined compression and unconfined compression moduli will be calculated from the equilibrium load cell readings, deformation, and measured cross-sectional area. Poisson's ratios will be calculated from digital images obtained by the optical system using a data reduction program that we have developed using MATLAB. Torsional shear modulus will be calculated from equilibrium measurements of torsional load, angular displacement, and diameter.*

This aim has been completed for the experimental data obtained during the 2004-2006 summer quarters.

Preliminary results were presented in [4].

**Methods.** The directions of the coordinate system used are related to a local anatomical coordinate system: 1=anteroposterior (AP), 2=mediolateral (ML), 3=axial (AX). Experimental data were analyzed from sequential confined compression (CC), unconfined compression (UCC), and torsional shear (TS) experiments on cylindrical bovine calf AC discs of mean diameter (d) of

3.2mm and height (h) of 1mm harvested in two orthogonal directions (h along ML and AX directions). UCC digital images were processed in Matlab to determine anisotropic Poisson's ratios  $\nu_{ij}$  (i=measured strain in direction of applied loading, j= measured strain in a transverse direction). Data from the CC and UCC experiments were used to determine strain-dependent functions for the confined compression modulus ( $H_A$ ) and Young's modulus (E).

**Results.** There was a significant difference found between  $\nu_{13}$  and  $\nu_{12}$ ,  $\nu_{31}$ , and  $\nu_{32}$  (Table 1). The stress-strain curves were nonlinear (data not shown). Permeability constants  $k_0$  and M were obtained using the strain-dependent function  $k = k_0 e^M$  (Table 2). In the TS experiments, actual superimposed compressive strains varied from 10%. Shear modulus ( $\mu$ ) was calculated by performing a linear regression between shear modulus and compressive stress ( $R^2 = 0.88$  for ML,  $R^2 = 0.99$  for AX) and determining the modulus at an offset stress corresponding to 10% strain (Table 2).

Table 1. Results (mean  $\pm$  S.D., n=8) measured in CC ( $H_A$  in MPa) and UCC (E in MPa,  $\nu_{ij}$ ). Subscripts 1, 2, and 3 refer to ML, AP, and AX directions, respectively. \* indicates significant difference ( $p < 0.05$ ).

	0% Strain	15% Strain	30% Strain	45% Strain
$H_{A1}$	0.438 $\pm$ 0.190	0.552 $\pm$ 0.229	0.699 $\pm$ 0.283	0.887 $\pm$ 0.358
$H_{A3}$	0.434 $\pm$ 0.143	0.543 $\pm$ 0.150	0.684 $\pm$ 0.167	0.866 $\pm$ 0.209
$E_1$	0.476 $\pm$ 0.213	0.549 $\pm$ 0.227	0.636 $\pm$ 0.252	0.741 $\pm$ 0.294
$E_3$	0.505 $\pm$ 0.147	0.616 $\pm$ 0.179	0.751 $\pm$ 0.221	0.918 $\pm$ 0.276
$\nu_{12}$	n/a	0.128 $\pm$ 0.063	0.129 $\pm$ 0.065	0.129 $\pm$ 0.090
$\nu_{13}$	n/a	0.212 $\pm$ 0.103 *	0.212 $\pm$ 0.078 *	0.225 $\pm$ 0.088 *
$\nu_{31}$	n/a	0.107 $\pm$ 0.043	0.112 $\pm$ 0.029	0.121 $\pm$ 0.023
$\nu_{32}$	n/a	0.103 $\pm$ 0.026	0.107 $\pm$ 0.030	0.109 $\pm$ 0.027

Table 2. Permeability constants  $k_0$ , M and shear modulus  $\mu$  measured in CC and TS experiments (mean  $\pm$  S.D., n=5). Subscripts 1 and 3 refer to ML and AX directions, respectively.

$k_{01}$ [ $10^{-16}$ m <sup>2</sup> /Pa·s]	5.892	$\pm$	2.681
$k_{03}$ [ $10^{-16}$ m <sup>2</sup> /Pa·s]	7.981	$\pm$	4.368
$M_1$	5.726	$\pm$	0.615
$M_3$	5.568	$\pm$	0.722
$\mu_1$ [MPa]	0.595	$\pm$	0.255
$\mu_3$ [MPa]	0.788	$\pm$	0.122

**Conclusions.** The results indicate that the protocols and the novel optical system enable the measurement of anisotropic material properties. The preliminary results (presented here) were used for completion of the remaining specific aims (see below). Final results are being incorporated into a manuscript that will be submitted to a peer-reviewed journal.

**Specific Aim 2:** A proteoglycan stress equation for finite deformations will be developed.

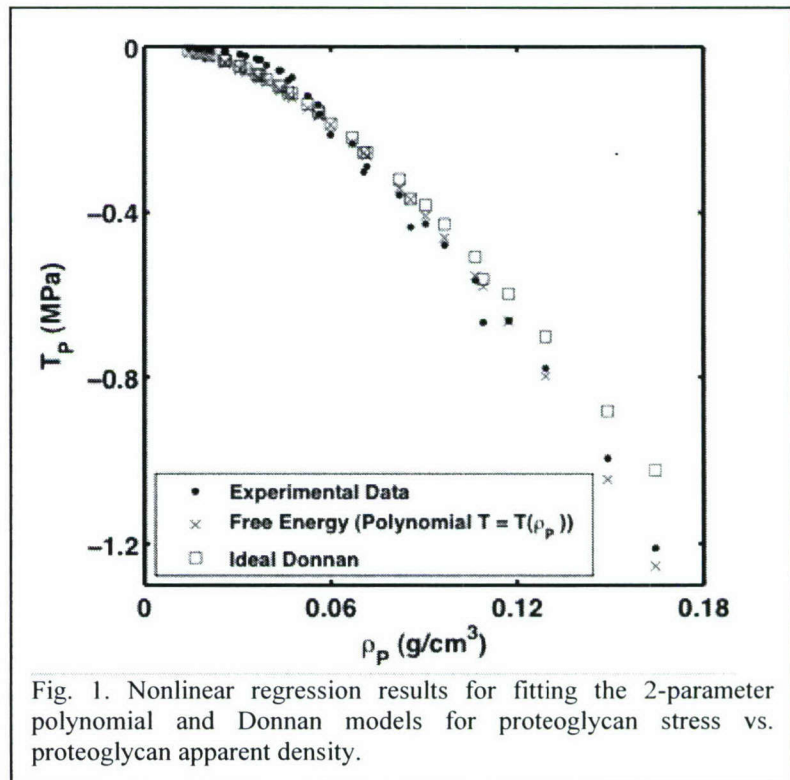
This aim has been completed using experimental results, including the data obtained during the 2004-2006 summers.

**Methods.** The purpose of this aim was to assess the accuracy of different proteoglycan (PG) continuum-level swelling stress models over the range of PG apparent density ( $\rho^p$ ) values expected through the full thickness of bovine calf AC explants. The material constants of several PG continuum-level swelling stress models were determined based on a published molecular-level model using published experimental data. The predictive capabilities of the models were assessed by computing  $R^2$  values resulting from the comparison of each individual model to the published molecular-level model.

In particular, Bassar et al. (Bassar, P. J., Schneiderman, R., Bank, R. A., Wachtel, E., and Maroudas, A., 1998, "Mechanical properties of the collagen network in human articular cartilage as measured by osmotic stress technique," *Arch Biochem Biophys*, 351, pp. 207-219) outlined a molecular-level analytical approach that is used to calculate an effective fixed charge density (FCD) for the PG component of biological tissues. The PGs are negatively charged molecules producing a FCD that accounts for the tissue's high compressive resistance. The PGs molecular size restricts them from the intrafibrillar (i.e., in the collagen (COL)) compartment and, therefore, reside solely in the extrafibrillar compartment. Here, the model of Bassar et al. is used to calculate the effective FCD and PG swelling stress using published experimental data that directly measured swelling stress vs. FCD for PG solutions (referred to below as "experimental data").

Initially, the dependence of continuum-level stress models on the individual water, COL, and PG apparent densities was tested by varying the experimental quantity of mass normalized by tissue wet weight, effectively varying apparent density, for each constituent independently over an experimental range of  $\pm 2$  standard deviations. PG swelling stress was seen to have a higher dependence on water than COL. From these results, four continuum-level stress models were selected for further analysis. Two forms depend only on PG density, while the other two depend on both PG and water density.

Results are only reported here for the 2-parameter polynomial form  $\mathbf{T}^p = -\alpha_1(\rho^p)^{\alpha_2}\mathbf{I}$ , where  $\mathbf{T}^p$  is the PG stress tensor,  $\rho^p$  is the PG apparent density (normalized by total tissue volume),  $\mathbf{I}$  is the identity tensor, and  $\alpha_1$  and  $\alpha_2$  are material constants. The predicted PG continuum-level swelling stress normalized to total tissue area ( $T^p$ ) is computed individually for each model for simulated confined compression tests of nine specimens from three different depths. The results using this stress equation are compared to those obtained using



the experimental model of Basser et al. and the ideal Donnan osmotic swelling pressure model. Nonlinear regression was performed in MATLAB (The MathWorks, V7.2), using the Levenberg-Marquardt algorithm.

1Results. The  $R^2$  values were 0.99 for the 2-parameter polynomial model and 0.97 for the ideal Donnan model (Fig. 1).

Conclusions. The results indicate that the 2-parameter PG continuum-level swelling stress model, dependent only on apparent PG density, will provide an excellent ( $R^2=0.99$ ) estimate of PG swelling stress for specimens typically found through the full articular thickness of bovine calf AC. Additionally, this work marks an important step from our previous models in that the equation depends on PG density, rather than the Jacobian of the PG deformation gradient tensor [1,5], eliminating dependence on the reference configuration and allowing application to specimens with different reference configuration densities. Preliminary results from this aim were used to accomplish the specific aim 3 (discussed below) and presented in [1,5]. Final results are being incorporated into an abstract that will be submitted to the 2007 ASME Summer Bioengineering Conference.

***Specific Aim 3:** We will determine a solid matrix stress-strain equation using the reduced form of the bimodular theory.*

Since the bimodular theory discussed in the 2005 proposal was not suitable for the finite element method [3], a new theory was developed [1,2,5].

Methods. The results presented here are based on a model that employs the stress balance hypothesis by using distinct stress constitutive equations for the PG and COL constituents. For the COL network, a strain energy function for finite deformations was developed [2] that models anisotropy using fiber families that behave in a bimodular fashion; i.e., the fibers only support tensile stresses. Two types of fiber families were introduced: primary fiber families that are oriented along anatomical directions defined relative to the split-line direction (direction 1), and secondary fiber families that represent crosslinking mechanisms between the primary fiber families. The proposed strain energy function satisfies numerical and material stability conditions, and contains 5 material constants and 3 angles that define the orientations of the secondary fibers. For the PGs, an isotropic strain energy function with 2 material constants was used [5].

Material constants were chosen based on experimental results (including results obtained during the 2004-2006 summers [4]) and the boundary-value problems of uniaxial tension (UT) and unconfined compression (UCC) were solved in three orthogonal directions.

Results. The results presented here (Fig. 2) [1,5] suggest that the main advantage of a model employing the secondary fiber families is to provide the capability for modeling anisotropic and asymmetric Poisson's ratios, as well as axial stress-axial strain responses, in UT and UCC for finite deformations. Furthermore, the model provides evidence in support of the stress balance hypothesis, as the predicted stress softening phenomena is similar to those recently measured.

Conclusions. This model has been employed in a validation study for our cartilage growth mixture (CGM) model [1]. The proposed CGM model was shown to match tissue biochemical content and volume while predicting theoretical values of tensile moduli that do not significantly differ from experimental measurements. Also, theoretical values of a scalar COL remodeling factor were positively correlated with COL crosslink density, and mass growth functions were positively correlated with cell density. The results provide baseline growth laws for free-swelling growth protocols for specific tissue sources.

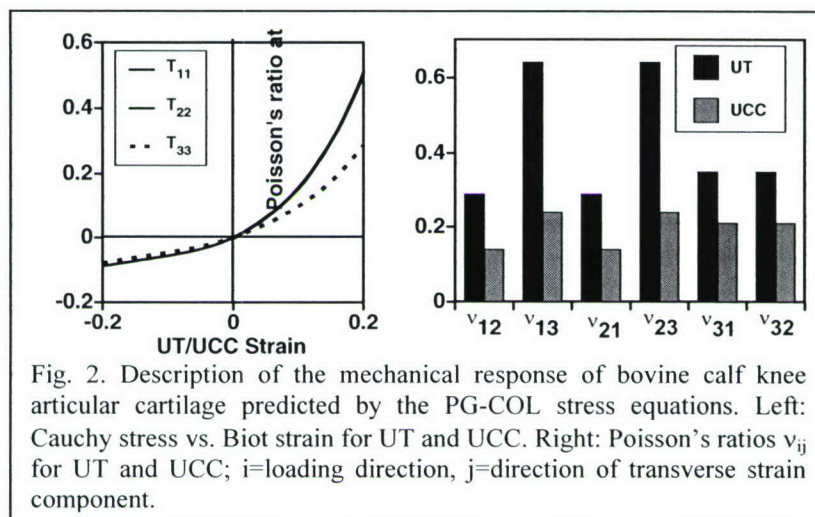


Fig. 2. Description of the mechanical response of bovine calf knee articular cartilage predicted by the PG-COL stress equations. Left: Cauchy stress vs. Biot strain for UT and UCC. Right: Poisson's ratios  $v_{ij}$  for UT and UCC;  $i$ =loading direction,  $j$ =direction of transverse strain component.

#### Publications.

1. Klisch SM, Asanbaeva A, Masuda K, Thonar EJ-M, Davol A, Sah RL. Refinement of a cartilage growth mixture model: polyconvex strain energy functions, collagen remodeling, and validation protocols. In preparation, 2006.
2. Klisch SM. Bimodular polyconvex fiber-reinforced strain energy functions for articular cartilage. *ASME Journal of Biomechanical Engineering*, in press, 2006.
3. Klisch SM. A bimodular theory for finite deformations: comparison of orthotropic second-order and exponential stress constitutive equations for articular cartilage. *Biomechanics and Modeling in Mechanobiology*, 5:90-101, 2006.
4. Ficklin TP, Thomas GC, Asanbaeva A, Chen AC, Sah RL, Davol A, Klisch SM. Development of an experimental protocol to measure anisotropic material properties of bovine articular cartilage. *Transactions of the ASME Summer Bioengineering Conference*, 2006.
5. Klisch SM, Sah RL, Davol A. Bimodular-orthotropic-polyconvex strain energy functions for the collagen-proteoglycan solid matrix of articular cartilage. *Transactions of the ASME Summer Bioengineering Conference*, 2006.

#### Full Publications Related to Project.

Copies of full publications follow.

# A Bimodular Polyconvex Anisotropic Strain Energy Function for Articular Cartilage

**Stephen M. Klisch**

Associate Professor  
Mechanical Engineering Department,  
California Polytechnic State University,  
San Luis Obispo, CA 93407  
e-mail: sklisch@calpoly.edu

*A strain energy function for finite deformations is developed that has the capability to describe the nonlinear, anisotropic, and asymmetric mechanical response that is typical of articular cartilage. In particular, the bimodular feature is employed by including strain energy terms that are only mechanically active when the corresponding fiber directions are in tension. Furthermore, the strain energy function is a polyconvex function of the deformation gradient tensor so that it meets material stability criteria. A novel feature of the model is the use of bimodular and polyconvex "strong interaction terms" for the strain invariants of orthotropic materials. Several regression analyses are performed using a hypothetical experimental dataset that captures the anisotropic and asymmetric behavior of articular cartilage. The results suggest that the main advantage of a model employing the strong interaction terms is to provide the capability for modeling anisotropic and asymmetric Poisson's ratios, as well as axial stress-axial strain responses, in tension and compression for finite deformations. [DOI: 10.1115/1.2486225]*

## Introduction

The extracellular solid matrix of articular cartilage contains proteoglycans and a crosslinked collagen network. The proteoglycans are negatively charged molecules that primarily resist compressive loads [1,2] while the collagen network primarily resists tensile and shear loads [3,4]. Due in part to its complex molecular structure, articular cartilage typically behaves as an anisotropic material with substantial tension-compression asymmetry [5–10] and likely experiences finite, multi-dimensional strains when subject to typical loads [11,12]. In particular, both the Young's modulus and Poisson's ratio<sup>1</sup> are anisotropic and strain dependent, and can be approximately two orders of magnitude greater in tension than in compression [9,10,13–19]. Consequently, the development of accurate finite deformation models of the equilibrium elastic response is challenging.

Bimodular elastic and biphasic models have been developed that can model the asymmetric tensile and compressive mechanical properties for infinitesimal strains [8,10]. Those models were based on a general bimodular theory for infinitesimal strains [20] in which the material constants may be discontinuous (or jump) across a surface of discontinuity in strain space, provided that stress continuity conditions are satisfied at the surface. Several exponential models for finite deformations allowing for different mechanical properties in tension and compression have been used for the arterial wall [21] and the annulus fibrosus [22]. However, those models have not employed a general bimodular theory that ensures stress continuity across the surface of discontinuity. Recently, a general bimodular theory employing second-order and exponential stress-strain equations was shown to be capable of modeling the anisotropy and asymmetry in Young's modulus for finite deformations [23]. Those results suggested that, when using the bimodular feature, second-order models might provide a material description as accurate as those provided by exponential models. However, the models studied in Ref. [23] were not ca-

pable of modeling the anisotropy and asymmetry in Poisson's ratio for finite deformations and were not appropriate for use in computational solutions.<sup>2</sup>

The overall goal of this study is to develop an elastic strain energy function for finite deformations of the articular cartilage solid matrix that meets several criteria. First, it should be capable of modeling the nonlinearity, anisotropy, and asymmetry in Young's modulus and Poisson's ratio. Although the desired accuracy of the stress-strain equation may not be the same for all applications, the level of accuracy sought here is likely to be crucial in continuum growth analysis.<sup>3</sup> In order to meet this criterion, the bimodular feature is employed. Second, it should satisfy stability criteria so that numerical stability of computational solutions can be expected. In order to meet this criterion, a polyconvex strain energy function is developed; polyconvexity guarantees the existence of local minimizers of the strain energy function when subject to boundary conditions [24] while not sharing the limitations of convexity with respect to the violation of invariance requirements and global uniqueness. Third, it should use a relatively low number of parameters needed to model the desired elastic response, so that the material constants are based on a model that is not over-parameterized<sup>4</sup> and can be determined from a combination of several common experimental protocols.

In a preliminary study, a bimodular polyconvex strain energy function was developed for articular cartilage based on the strain invariants for an orthotropic material [25]; however, that model, nor the earlier second-order and exponential models [23], were capable of modeling the anisotropy and asymmetry of Poisson's ratio. In that polyconvex model [25], there were no strong interaction or coupling terms for the orthotropic strain invariants. Although recent studies have proposed [26] or used [27] strong interaction terms for orthotropic strain invariants, preliminary studies for this work were not successful in using those terms with the bimodular feature.<sup>5</sup> Also, those studies have not discussed

<sup>1</sup>In this paper, the terms "Young's modulus" and "Poisson's ratio" will be used to refer to strain-dependent functions because a finite deformation theory is used.

Contributed by the Bioengineering Division of ASME for publication in the JOURNAL OF BIOMECHANICAL ENGINEERING. Manuscript received May 1, 2006; final manuscript received September 15, 2006. Review conducted by Clark T. Hung.

<sup>2</sup>The second-order model in terms of the first Piola-Kirchhoff stress developed in Ref. [23] was shown to satisfy stability criteria; however, the corresponding Cauchy stress was not.

<sup>3</sup>Applications are presented in the "Discussion."

<sup>4</sup>See the "Discussion" for comments on "over-parameterization" in the context of the nonlinear regression analysis used here.

<sup>5</sup>In particular, strong interaction terms that satisfy the bimodular stress-strain continuity conditions stated in Eq. (7) were not found for an orthotropic material.

possible microstructural causes. An aim of this work is to derive bimodular strong interaction terms that are “simple” enough to allow the experimenter to investigate possible microstructural mechanisms.

A recent development of Ref. [28] used two mechanically equivalent secondary fiber families, in addition to primary fiber families, to represent the phenomena of collagen crosslinking in the annulus fibrosus tissue; that model was capable of producing tensile Poisson’s ratios that are an order of magnitude greater than those of our earlier studies [23,25]. Since the secondary fiber families introduced in Ref. [28] basically serve as strong interaction or coupling terms for the strain invariants related to the primary fiber families, here it was hypothesized that the introduction of strain invariants generated by secondary fiber families will allow a more accurate description of tensile Poisson’s ratios for articular cartilage. In contrast to Ref. [28], this development is incorporated into a bimodular polyconvex strain energy function.

The specific objectives are to: (1) adapt the bimodular theory for finite deformations to the present application; (2) develop a bimodular polyconvex anisotropic strain energy function using primary fibers and strong interaction terms generated by secondary fibers; and (3) compare the predictive capability of models with and without the strong interaction terms using experimental data gathered from the literature. The results suggest that using both the bimodular feature and the strong interaction terms facilitates the accurate description of the anisotropic and asymmetric mechanical properties of articular cartilage in finite deformations.

## Methods

**Background.** The right Cauchy–Green deformation tensor  $\mathbf{C}$  is defined as

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} \quad (1)$$

where  $\mathbf{F}$  is the deformation gradient tensor and the superscript  $T$  signifies the transpose operator. The Cauchy, first Piola–Kirchhoff, and second Piola–Kirchhoff stress tensors (denoted as  $\mathbf{T}$ ,  $\mathbf{P}$ , and  $\mathbf{S}$ , respectively) are related by

$$\mathbf{J} \mathbf{T} = \mathbf{P} \mathbf{F}^T = \mathbf{F} \mathbf{S}^T \quad (2)$$

where  $J$  is the determinant of  $\mathbf{F}$ . The stress constitutive equations for a Green-elastic material may be expressed as

$$\mathbf{S} = 2 \frac{\partial W}{\partial \mathbf{C}} = 2 \sum_{i=1}^n \frac{\partial W}{\partial I_i} \frac{\partial I_i}{\partial \mathbf{C}} \quad (3)$$

where  $W = \hat{W}\{I_i\}$  is a scalar strain energy function that depends on a set of invariants  $I_i$  corresponding to the material symmetry group. The fourth-order elasticity tensor is defined as

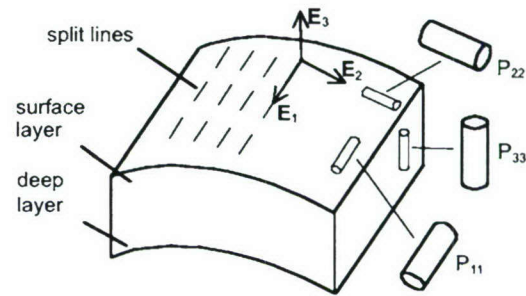
$$\mathbf{C} = \frac{\partial \mathbf{S}}{\partial \mathbf{C}} \quad (4)$$

**Bimodular Elasticity for Finite Deformations.** Due to the observed tension–compression asymmetry of the articular cartilage solid matrix, a bimodular theory is used. Earlier models [8,10,23] were based on a bimodular theory [20] in which the material constants may be discontinuous (or jump) across a surface of discontinuity in strain space, provided that stress continuity conditions are satisfied at the surface. The bimodular theory of Ref. [20] was developed in terms of the second Piola–Kirchhoff stress and Lagrange strain tensors. Here, that theory is reformulated to use  $\mathbf{C}$  instead of the Lagrange strain tensor.

A scalar valued function of  $\mathbf{C}$  identifying a surface of discontinuity in the six-dimensional strain space of  $\mathbf{C}$  is defined as

$$g(\mathbf{C}) = 0 \quad (5)$$

and is restricted to be a function of the invariants corresponding to the material symmetry group. Different strain energy functions may be specified on either side of a surface of discontinuity; i.e.



**Fig. 1 Schematic of the coordinate system and experimental specimen orientations in relation to anatomical directions. The unit vector  $\mathbf{E}_1$  is parallel to the local split-line direction, the unit vector  $\mathbf{E}_3$  is perpendicular to the articular surface, and the unit vector  $\mathbf{E}_2$  is perpendicular to the split-line direction and parallel to the surface. The cylinders labeled  $P_{11}$ ,  $P_{22}$ , and  $P_{33}$  represent specimens loaded in tension or compression along the  $\mathbf{E}_1$ ,  $\mathbf{E}_2$ , and  $\mathbf{E}_3$  directions, respectively.**

$$W = W_+ \text{ if } g(\mathbf{C}) > 0, \quad W = W_- \text{ if } g(\mathbf{C}) < 0 \quad (6)$$

In a similar fashion, different stress and elasticity tensors may be specified on either side of a surface of discontinuity; i.e., as  $(\mathbf{S}_+, \mathbf{S}_-, \mathbf{C}_+, \mathbf{C}_-)$

In Ref. [20], a theorem was proved establishing necessary and sufficient conditions for stress continuity across the surface of discontinuity. Introducing a slight modification in Lemma 3.2 of Ref. [20], one obtains the following necessary and sufficient conditions for stress continuity across the surface of discontinuity

$$\mathbf{S} = \mathbf{S}_+ = \mathbf{S}_-, \quad [[\mathbf{C}]] = \mathbf{C}_+ - \mathbf{C}_- = s(\mathbf{C}) \frac{\partial g}{\partial \mathbf{C}} \otimes \frac{\partial g}{\partial \mathbf{C}} \quad (7)$$

for all  $\mathbf{C}$  that satisfy  $g(\mathbf{C})=0$ , where  $[[\mathbf{C}]]$  represents the jump in the elasticity tensor,  $s(\mathbf{C})$  is a scalar valued function of  $\mathbf{C}$ , and  $\otimes$  is the tensor dyadic product.

**Structural anisotropy.** Spencer [29] proposed a general theory capable of modeling an anisotropic material as a composite material consisting of an isotropic matrix reinforced with fiber families. That theory has been used to develop strain energy functions for cartilaginous tissues [22,28,30–32]. For example, in Ref. [30] two mechanically equivalent fiber families were used to model the annulus fibrosus in finite deformations; in Ref. [28] that model was generalized to include two mechanically equivalent fiber families representing crosslinking phenomena. Also, in Ref. [21] two families of fibers were used to model arterial tissue, but these fiber families were not assumed to be mechanically equivalent because a bimodular feature was used.<sup>6</sup> Here, secondary fibers are used (as in Ref. [28]) without assuming that the fiber families are mechanically equivalent so that the bimodular feature can be used (as in Ref. [21]).

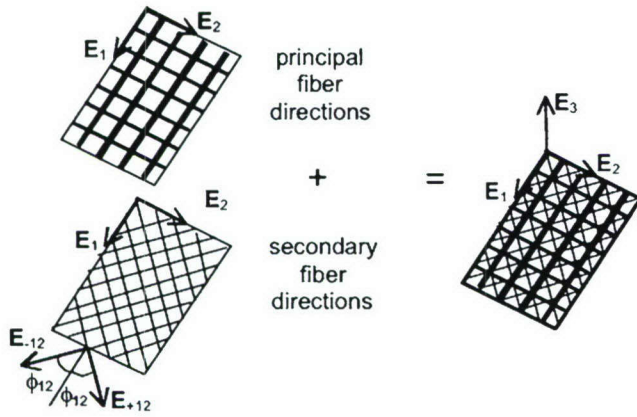
First, three fiber families are introduced that are parallel to three mutually orthogonal basis vectors ( $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ ) in a stress-free reference configuration; these will be referred to as “principal fibers.” As seen below in Eq. (13), the principal fibers generate strain invariants for orthotropic materials. Structural tensors ( $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3$ ) are defined as in the case of an orthotropic material

$$\mathbf{M}_1 = \mathbf{E}_1 \otimes \mathbf{E}_1, \quad \mathbf{M}_2 = \mathbf{E}_2 \otimes \mathbf{E}_2, \quad \mathbf{M}_3 = \mathbf{E}_3 \otimes \mathbf{E}_3 \quad (8)$$

The unit vectors used to form these structural tensors correspond to the following anatomical directions:  $\mathbf{E}_1$  is parallel to the local split-line direction,  $\mathbf{E}_3$  is perpendicular to the articular surface, and  $\mathbf{E}_2$  is perpendicular to the split-line direction and parallel to the surface (Fig. 1).

Second, two fiber families are introduced in each of the three

<sup>6</sup>See the comment below following Eq. (21).



**Fig. 2 Schematic of the principal and secondary fiber orientations in relation to anatomical directions in the 1-2 plane. The two principal fiber directions are parallel to the unit vectors  $\mathbf{E}_1$  and  $\mathbf{E}_2$  and the two secondary fiber directions, denoted as  $\mathbf{E}_{\pm 12}$ , are oriented at angles of  $\pm\phi_{12}$  to the  $\mathbf{E}_1$  direction. The weights of the line elements represent the relative strength of the fiber directions as predicted by regression analysis; i.e., the principal fibers along the  $\mathbf{E}_1$  direction are the strongest while the secondary fibers are the weakest.**

planes formed by the basis vectors ( $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ ); these will be referred to as “secondary fibers.” As seen below in Eq. (14), the secondary fibers generate strain invariants that represent strong interaction terms between the orthotropic strain invariants. Consider the 1-2 plane, which contains the ( $\mathbf{E}_1, \mathbf{E}_2$ ) unit vectors (Fig. 2). The secondary fiber families are defined to lie at angles of  $\pm\phi_{12}$  to the  $\mathbf{E}_1$  direction. These secondary fiber directions are denoted as ( $\mathbf{E}_{+12}, \mathbf{E}_{-12}$ ) and are expressed as

$$\mathbf{E}_{\pm 12} = \cos \phi_{12} \mathbf{E}_1 \pm \sin \phi_{12} \mathbf{E}_2 \quad (9)$$

Corresponding structural tensors ( $\mathbf{M}_{+12}, \mathbf{M}_{-12}$ ) are defined as in Eq. (8)

$$\begin{aligned} \mathbf{M}_{\pm 12} = \mathbf{E}_{\pm 12} \otimes \mathbf{E}_{\pm 12} = & \cos^2 \phi_{12} \mathbf{E}_1 \otimes \mathbf{E}_1 + \sin^2 \phi_{12} \mathbf{E}_2 \\ & \otimes \mathbf{E}_2 \pm \cos \phi_{12} \sin \phi_{12} (\mathbf{E}_1 \otimes \mathbf{E}_2 + \mathbf{E}_2 \otimes \mathbf{E}_1) \end{aligned} \quad (10)$$

The secondary fiber directions introduced in the 1-3 and 2-3 planes are denoted as ( $\mathbf{E}_{+13}, \mathbf{E}_{-13}$ ) and ( $\mathbf{E}_{+23}, \mathbf{E}_{-23}$ ), respectively, are expressed as

$$\mathbf{E}_{\pm 13} = \cos \phi_{13} \mathbf{E}_1 \pm \sin \phi_{13} \mathbf{E}_3, \quad \mathbf{E}_{\pm 23} = \cos \phi_{23} \mathbf{E}_2 \pm \sin \phi_{23} \mathbf{E}_3 \quad (11)$$

Corresponding structural tensors ( $\mathbf{M}_{+13}, \mathbf{M}_{-13}, \mathbf{M}_{+23}, \mathbf{M}_{-23}$ ) are defined as in Eq. (10)

$$\begin{aligned} \mathbf{M}_{\pm 13} = & \cos^2 \phi_{13} \mathbf{E}_1 \otimes \mathbf{E}_1 + \sin^2 \phi_{13} \mathbf{E}_3 \otimes \mathbf{E}_3 \\ & \pm \cos \phi_{13} \sin \phi_{13} (\mathbf{E}_1 \otimes \mathbf{E}_3 + \mathbf{E}_3 \otimes \mathbf{E}_1) \\ \mathbf{M}_{\pm 23} = & \cos^2 \phi_{23} \mathbf{E}_2 \otimes \mathbf{E}_2 + \sin^2 \phi_{23} \mathbf{E}_3 \otimes \mathbf{E}_3 \\ & \pm \cos \phi_{23} \sin \phi_{23} (\mathbf{E}_2 \otimes \mathbf{E}_3 + \mathbf{E}_3 \otimes \mathbf{E}_2) \end{aligned} \quad (12)$$

Following Ref. [29], the strain energy function  $W$  is assumed to be an isotropic function of  $\mathbf{C}$  and the nine structural tensors introduced above. In Ref. [33], a procedure is outlined for obtaining minimal lists of irreducible scalar invariants for an arbitrary finite number of symmetric structural tensors; however, that procedure was only employed for up to six symmetric structural tensors. Here, only the decoupled first-order scalar invariants are used in an attempt to obtain a relatively low number of material constants to prevent the model from becoming overparameterized. Invariants associated with the primary fibers include

$$\{\mathbf{M}_1 \cdot \mathbf{C}, \mathbf{M}_2 \cdot \mathbf{C}, \mathbf{M}_3 \cdot \mathbf{C}\} = \{C_{11}, C_{22}, C_{33}\} = \{\lambda_1^2, \lambda_2^2, \lambda_3^2\} \quad (13)$$

where  $\lambda_A^2$  represents the square of the stretch of the material line element initially oriented along the principal fiber direction  $\mathbf{E}_A$ . Invariants associated with the secondary fibers include

$$\begin{aligned} \{\mathbf{M}_{\pm 12} \cdot \mathbf{C}, \mathbf{M}_{\pm 13} \cdot \mathbf{C}, \mathbf{M}_{\pm 23} \cdot \mathbf{C}\} \\ = \{C_{11} \cos^2 \phi_{12} + C_{22} \sin^2 \phi_{12} \\ \pm 2C_{12} \cos \phi_{12} \sin \phi_{12}, C_{11} \cos^2 \phi_{13} \\ + C_{33} \sin^2 \phi_{13} \pm 2C_{13} \cos \phi_{13} \sin \phi_{13}, C_{22} \cos^2 \phi_{23} \\ + C_{33} \sin^2 \phi_{23} \pm 2C_{23} \cos \phi_{23} \sin \phi_{23}\} \\ = \{\lambda_{\pm 12}^2, \lambda_{\pm 13}^2, \lambda_{\pm 23}^2\} \end{aligned} \quad (14)$$

where  $\lambda_{+AB}^2$  and  $\lambda_{-AB}^2$  represent the squares of the stretches of the material line elements initially oriented along the secondary fiber directions  $\mathbf{E}_{+AB}$  and  $\mathbf{E}_{-AB}$ , respectively. Note that these invariants represent strong interaction or coupling terms for the orthotropic strain invariants as discussed in Refs. [26,34]; for example, the invariant  $\lambda_{+12}^2$  is a function of the invariants  $C_{11} = \lambda_1^2$  and  $C_{22} = \lambda_2^2$ , thereby coupling these invariants. It is important to emphasize that this approach adopts a lesser level of material symmetry than orthotropy because the invariants in Eq. (14) can easily be shown to not be invariant under transformations due to reflections about three orthogonal planes.<sup>7</sup>

Including the three principal invariants of  $\mathbf{C}$ , the strain energy function for the model proposed here can be expressed as a function of 12 invariants

$$W = \hat{W}\{\text{tr } \mathbf{C}, \text{tr}(\text{adj } \mathbf{C}), \det \mathbf{C}, \lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_{\pm 12}^2, \lambda_{\pm 13}^2, \lambda_{\pm 23}^2\} \quad (15)$$

where  $\text{tr}$  is the trace operator,  $\text{adj } \mathbf{C} = (\det \mathbf{C})\mathbf{C}^{-1}$  is the adjugate of  $\mathbf{C}$ , and  $\det$  is the determinant operator.

**Bimodular Polyconvex Strain Energy functions.** In recent years, polyconvex strain energy functions have been proposed for anisotropic materials [26,27,34,36]; discussion of the rationale for using polyconvex strain energy functions is in the “Introduction.” A sufficient condition for polyconvexity is as follows [26]: if the strain energy function  $W(\mathbf{F})$  satisfies the additive decomposition

$$W(\mathbf{F}) = W_1(\mathbf{F}) + W_2(\text{adj } \mathbf{F}) + W_3(\det \mathbf{F}) \quad (16)$$

and each of the functions ( $W_1(\mathbf{F}), W_2(\text{adj } \mathbf{F}), W_3(\det \mathbf{F})$ ) is a convex function of ( $\mathbf{F}, \text{adj } \mathbf{F}, \det \mathbf{F}$ ), respectively, then  $W(\mathbf{F})$  is polyconvex. Furthermore, addition of two or more polyconvex functions results in a polyconvex function.

Here,  $W$  is additively decomposed into two terms  $W_O$  and  $W_{\text{BIM}}$  representing nonbimodular and bimodular contributions, respectively. In general,  $W_O$  can be anisotropic; a general polynomial form is proposed in Ref. [36]. Here, a simple isotropic function is adopted from Ref. [36] for  $W_O$

$$W_O = \frac{1}{2}\mu\{(\text{tr } \mathbf{C} - 3) + (\text{tr}[\text{adj } \mathbf{C}] - 3) - 3 \ln(\det \mathbf{C})\} \quad (17)$$

This term is polyconvex if  $\mu$  is positive and contributes a stress term as follows [36]

$$\mathbf{S}_O = \mu\{\mathbf{I} - (\det \mathbf{C})\mathbf{C}^{-2} + [(\det \mathbf{C})\text{tr } \mathbf{C}^{-1} - 3]\mathbf{C}^{-1}\} \quad (18)$$

Then, it is assumed that  $W_{\text{BIM}}$  represents the collagen network molecules that account for all of the tissue anisotropy. To model tension-compression asymmetry, it is assumed that all fiber families can only support tensile stresses; consequently, a total of nine surfaces of discontinuity are used

<sup>7</sup>If the secondary fibers are not bimodular and assumed mechanically equivalent, then the symmetry reduces to orthotropy as in [30,35].

$$g_1 = \mathbf{M}_1 \cdot \mathbf{C} - 1 = 0, \quad g_2 = \mathbf{M}_2 \cdot \mathbf{C} - 1 = 0, \quad g_3 = \mathbf{M}_3 \cdot \mathbf{C} - 1 = 0$$

$$g_{\pm 12} = \mathbf{M}_{\pm 12} \cdot \mathbf{C} - 1 = 0, \quad g_{\pm 13} = \mathbf{M}_{\pm 13} \cdot \mathbf{C} - 1 = 0$$

$$g_{\pm 23} = \mathbf{M}_{\pm 23} \cdot \mathbf{C} - 1 = 0 \quad (19)$$

For example, the surface  $g_1 = \mathbf{M}_1 \cdot \mathbf{C} - 1 = \lambda_1^2 - 1 = 0$  defines a five-dimensional hyperplane that divides the  $\mathbf{C}$  space into two half-spaces corresponding to tensile and compressive strains in the principal fiber direction  $\mathbf{E}_1$ . The following bimodular form is used

$$W_{\text{BIM}} = \frac{1}{6} \gamma_1 [\lambda_1] (\lambda_1^2 - 1)^3 + \frac{1}{6} \gamma_2 [\lambda_2] (\lambda_2^2 - 1)^3 + \frac{1}{6} \gamma_3 [\lambda_3] (\lambda_3^2 - 1)^3 \\ + \frac{1}{6} \delta_{\pm 12} [\lambda_{\pm 12}] (\lambda_{\pm 12}^2 - 1)^3 + \frac{1}{6} \delta_{\pm 13} [\lambda_{\pm 13}] (\lambda_{\pm 13}^2 - 1)^3 \\ + \frac{1}{6} \delta_{\pm 23} [\lambda_{\pm 23}] (\lambda_{\pm 23}^2 - 1)^3 \quad (20)$$

where  $(\gamma_1, \gamma_2, \gamma_3, \delta_{\pm 12}, \delta_{\pm 13}, \delta_{\pm 23})$  are six material constants that represent bimodular terms via the definitions

$$\gamma_1 [\lambda_1] = \begin{cases} \gamma_1 > 0 & \text{if } \lambda_1 > 1 \\ 0 & \text{if } \lambda_1 < 1 \end{cases} \\ \delta_{\pm 12} [\lambda_{\pm 12}] = \begin{cases} \delta_{\pm 12} > 0 & \text{if } \lambda_{\pm 12} > 1 \\ 0 & \text{if } \lambda_{\pm 12} < 1 \end{cases}, \quad \text{etc} \quad (21)$$

and the angles  $(\phi_{12}, \phi_{13}, \phi_{23})$  that appear in Eq. (14) can be regarded as three additional material constants. In this general formulation, the two secondary fiber families in any of the three planes will have the same stiffness if both are active (i.e.,  $\delta_{\pm 12} = \delta_{-12}$ ), but are not assumed to be mechanically equivalent as defined in Ref. [29] because in some shearing deformations one fiber family may be in tension while the other may be in compression. Considering Eq. (21), it is evident that each of the material constants are related to one of the surfaces of discontinuity defined in Eq. (19). For example, the material constant  $\gamma_1$  defines a strain energy term that can jump across the surface  $g_1 = \mathbf{M}_1 \cdot \mathbf{C} - 1 = \lambda_1 - 1 = 0$ .

In the Appendix, the proposed strain-energy function  $W_{\text{BIM}}$  is shown to satisfy both the bimodular stress continuity and polyconvexity conditions when the material constants are defined as in Eq. (21). Since preliminary statistical results suggested that the model defined by Eq. (20) was overparameterized given the experimental dataset,<sup>8</sup> it is further assumed that the material constants associated with the secondary fibers are equal (when active); i.e.,  $\delta_{\pm AB} = \delta$ . This reduced strain energy function contributes a stress term as

$$\mathbf{S}_{\text{BIM}} = \gamma_1 [\lambda_1] (\lambda_1^2 - 1)^2 \mathbf{M}_1 + \gamma_2 [\lambda_2] (\lambda_2^2 - 1)^2 \mathbf{M}_2 + \gamma_3 [\lambda_3] (\lambda_3^2 - 1)^2 \mathbf{M}_3 \\ + \delta [\lambda_{\pm 12}] (\lambda_{\pm 12}^2 - 1)^2 \mathbf{M}_{\pm 12} + \delta [\lambda_{\pm 13}] (\lambda_{\pm 13}^2 - 1)^2 \mathbf{M}_{\pm 13} \\ + \delta [\lambda_{\pm 23}] (\lambda_{\pm 23}^2 - 1)^2 \mathbf{M}_{\pm 23} \quad (22)$$

so that the stress constitutive equation is defined by Eqs. (18) and (22) as

$$\mathbf{S} = \mathbf{S}_O + \mathbf{S}_{\text{BIM}} \quad (23)$$

with a total of eight material constants  $(\mu, \gamma_1, \gamma_2, \gamma_3, \delta, \phi_{12}, \phi_{13}, \phi_{23})$ .

**Experimental Data.** A hypothetical experimental dataset was developed that approximates the equilibrium elastic response of the solid matrix of adult human cartilage in the surface region (Tables 1 and 2), assuming homogeneous tissue composition and elastic properties of test specimens that are free of residual stress. In order to construct enough data to prevent the models presented here from being overparameterized, it was necessary to use data

**Table 1 Values of tangent Young's modulus (MPa) in tension at 0% strain ( $E_{+0}$ ) and 16% strain ( $E_{+0.16}$ ) and in compression at 0% strain ( $E_{-0}$ ) and 16% strain ( $E_{-0.16}$ ) in the 1, 2, and 3 directions for the experimental dataset used.**

Parameter	Direction		
	1	2	3
$E_{+0}$	7.8	5.9	1.2
$E_{+0.16}$	42.8	26.3	9.0
$E_{-0}$	0.18	0.18	0.18
$E_{-0.16}$	0.26	0.26	0.26

from several studies representing different anatomic locations, species, etc.

Based on how mechanical properties were calculated in the studies used here, the first Piola–Kirchhoff stress and Biot strain tensors are used. In particular, the first Piola–Kirchhoff stress normalizes load by original cross-sectional area. Also, the Biot strain tensor has principal strain components (e.g.,  $E_{11}^B, E_{22}^B, E_{33}^B$ ) that correspond to the definition of the infinitesimal strain tensor  $\boldsymbol{\epsilon}$  (e.g.,  $E_{11}^B = \epsilon_{11} = \lambda_1 - 1$ , etc.). Consequently, Poisson's ratios defined in terms of the Biot strain tensor correspond to the Poisson's ratios defined in terms of  $\boldsymbol{\epsilon}$  used in the studies mentioned below. For example, the Poisson's ratio  $\nu_{12}$  is defined here as

$$\nu_{12} = -E_{22}^B / E_{11}^B = -(\lambda_2 - 1) / (\lambda_1 - 1) = -\epsilon_{22} / \epsilon_{11} \quad (24)$$

The data used corresponds to uniaxial tension (UT) and unconfined compression (UCC) experiments along three directions: 1=parallel to the split-line; 2=perpendicular to the split line and parallel to the surface; and 3=perpendicular to the surface (Fig. 1). Exponential functions were used to generate axial (i.e., along the direction of applied loading) stress–axial strain data and linear functions were used to generate transverse strain–axial strain data from 0% to 20% strain in 2% increments. UT axial stress–axial strain data were adopted from Refs. [17,18]. UT Poisson's ratios were assumed based on the results of several studies [13,15,16,18,37]. UCC axial stress–axial strain data were adopted from Refs. [9,14,18] and assumed to be the same in all three directions.<sup>9</sup> UCC Poisson's ratios were assumed based on the results of Refs. [10,19,38,39].

It is important to note that this hypothetical dataset includes not only substantial anisotropy and asymmetry in the axial stress–strain response (Table 1), but also substantial anisotropy and asymmetry in the Poisson's ratios (Table 2). In particular, the Poisson's ratios in UT can be approximately two orders of magnitude greater than those in UCC, and in both UT and UCC the Poisson's ratios  $\nu_{13}$  and  $\nu_{23}$  have been measured to be greater than those in other directions; see Refs. [13,37] for UT and Refs. [10,19] for UCC.

**Regression Analysis.** A simultaneous nonlinear regression algorithm was performed in Mathematica (Wolfram, V5.0) based on an approach developed in Refs. [23,30]. The Levenberg–Marquardt method is used to minimize an error term representing the sum of squared differences between theoretical and experimental stress values. Although additional models were studied,<sup>10</sup> only results of three regression analyses are presented (Table 3). An eight-parameter model (8-PAR) defined by Eqs. (18) and (22) was studied. To provide a comparison with a model that does not use the strong interaction terms, a four-parameter model (4-PAR) was studied, for which the material constant  $\delta$  was set equal to zero. For the 8- and 4-PAR models, the assumed Poisson's ratios were used to prescribe the transverse strains for UT and UCC.

<sup>8</sup>These results are summarized in the “Discussion.”

<sup>9</sup>This limitation is addressed in the “Discussion.”

<sup>10</sup>The results of other models are summarized in the Discussion.

**Table 2 Numerical values of Poisson's ratios in tension ( $\nu_{+ij}$ ) and compression ( $\nu_{-ij}$ ) obtained from regression analysis;  $i$ =loading direction,  $j$ =direction of transverse strain component. The range given corresponds to values at 0% and 20% strain. The assumed values are given for comparison in the column labeled Range; tensile values were assumed to be linear functions of strain and the compressive values were assumed to be constant.**

Parameter	Range	Model		
		8-PAR	4-PAR	8-PAR-B
$\nu_{+12}$	0.5→1.0	0.44→0.83	0.25→0.26	0.79→1.21
$\nu_{+13}$	1.0→2.0	0.75→1.59	0.25→0.26	1.22→1.73
$\nu_{+21}$	0.5→1.0	0.46→0.86	0.25→0.26	0.96→1.54
$\nu_{+23}$	1.0→2.0	0.75→1.59	0.25→0.26	1.08→1.82
$\nu_{+31}$	0.5→1.0	0.32→0.45	0.25→0.26	0.63→0.81
$\nu_{+32}$	0.5→1.0	0.32→0.45	0.25→0.26	0.71→0.92
$\nu_{-12}$	0.1	0.10→0.03	0.11→0.06	0.01→-0.06
$\nu_{-13}$	0.2	0.19→0.09	0.60→0.10	0.08→0.04
$\nu_{-21}$	0.1	0.09→0.03	0.12→-0.06	0.01→-0.06
$\nu_{-23}$	0.2	0.20→0.09	0.21→0.11	0.12→0.12
$\nu_{-31}$	0.1	0.11→0.04	0.13→0.05	0.05→0.02
$\nu_{-32}$	0.1	0.13→0.05	0.15→0.06	0.03→0.01

Then, a composite function representing a total of 18 equations was derived: six axial stress–axial strain equations (three each in UT and UCC), and 12 transverse stress–axial strain equations corresponding to the traction-free boundary conditions (six each in UT and UCC). To provide a comparison with a model that does not explicitly include the 12 traction-free boundary condition equations, an additional regression with the eight-parameter model was performed that included only the six axial stress–axial strain equations obtained after prescribing the transverse strains (8-PAR-B). In all cases, the UCC stress values were weighted by multiplying each stress value by 100, since the UT stress response is two orders of magnitude greater than the UCC stress response.

After the nonlinear regression analysis was performed, the determined model parameters were used to derive numerical solutions to the UT and UCC boundary-value problems, including theoretical predictions of Poisson's ratios.

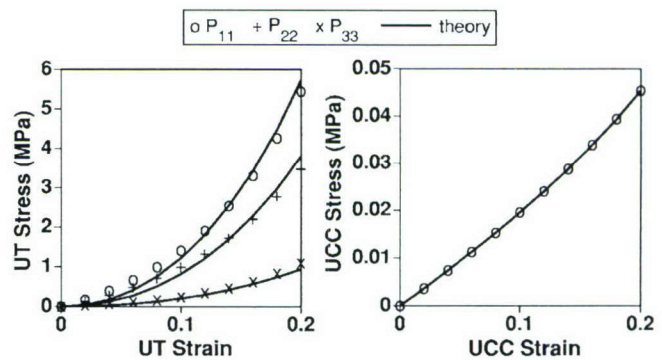
## Results

The numerical values for the material constants are presented in Table 3. The nonlinear regression analyses always converged to results consistent with the stability criteria; i.e., ( $\mu$ ,  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ,  $\delta$ ) were all positive. The calculated error terms were 0.773 and 0.876 for the 8-PAR and 4-PAR models, respectively, and 0.206 for the 8-PAR-B model. It is important to note that this latter error term cannot be directly compared to the others because fewer equations were used in the nonlinear regression.

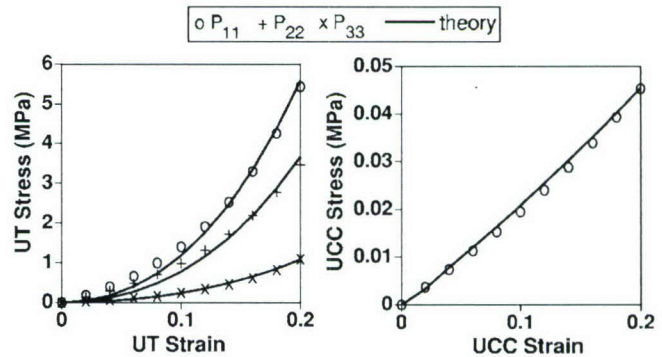
**Table 3 Numerical values for the material parameters obtained from regression analysis. The constants ( $\mu$ ,  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ,  $\delta$ ) are in MPa and the constants ( $\phi_{12}$ ,  $\phi_{13}$ ,  $\phi_{23}$ ) are in degrees.**

Parameter	Model		
	8-PAR	4-PAR	8-PAR-B
$\mu$	0.035	0.035	0.035
$\gamma_1$	23.13	23.88	21.92
$\gamma_2$	14.83	15.61	1.78
$\gamma_3$	3.86	4.51	3.26
$\delta$	12.07	—	436.8
$\phi_{12}$	45	—	46
$\phi_{13}$	35	—	40
$\phi_{23}$	35	—	42

The predictions of the 8-PAR and 4-PAR models (Figs. 3 and 4), as well as the 8-PAR-B model for axial stresses were qualitatively similar, with one exception. For the 8-PAR-B model that did not explicitly include the traction-free boundary conditions,



**Fig. 3 Predictions of the eight-parameter model (8-PAR) for the uniaxial tension (UT) response in the 1, 2, and 3 directions and the unconfined compression (UCC) response in the 1 direction. The theoretical UCC curves in the 2 and 3 directions are within 1% of the curve shown. UCC stress and strain values, although negative by definition, are plotted as positive numbers.**



**Fig. 4 Predictions of the four-parameter model (4-PAR) for the uniaxial tension (UT) response in the 1, 2, and 3 directions and the unconfined compression (UCC) response in the 1 direction. The theoretical UCC curves in the 2 and 3 directions are within 3% of the curve shown. UCC stress and strain values, although negative by definition, are plotted as positive numbers.**

the discrepancy between the assumed and theoretical values for UT in the 2 direction was relatively large. Comparing the 8-PAR and 8-PAR-B model results, it is evident that excluding the traction-free boundary condition equations in the regression analysis results in different predicted material constants (Table 3). In particular, the predicted strength  $\delta$  of the strong interaction terms was different by an order of magnitude (i.e., 12.07 for the 8-PAR model, 436.8 for the 8-PAR-B model).

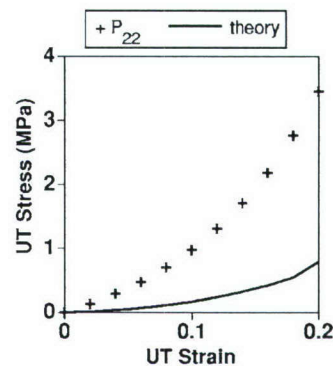
In contrast, predictions for Poisson's ratios varied substantially among the models (Table 2). Comparing the 8-PAR and 4-PAR model results, it is evident that including the strong interaction terms facilitates modeling the anisotropic and asymmetric Poisson's ratios, allowing the Poisson's ratios in UT to be 1–2 orders of magnitude greater than those in UCC. Also, the 8-PAR-B model, as compared to the 8-PAR model, provided predicted Poisson's ratios in better and worse agreement with the assumed values in UT and UCC, respectively.

## Discussion

In this paper, a bimodular polyconvex anisotropic strain energy function was developed with the aim of accurately modeling the anisotropic and asymmetric mechanical properties of articular cartilage. The 4-PAR model, based on a bimodular orthotropic material without strong interaction terms, is capable of providing reasonable predictions of the assumed axial stress–axial strain properties in three anatomically relevant directions. However, that model provides a poor description of anisotropic and asymmetric Poisson's ratios. In contrast, the 8-PAR model, which included strong interactions terms for the orthotropic strain invariants, provides a reasonable prediction of anisotropic and asymmetric Poisson's ratios as well as axial stress–axial strain properties. Also, the 8-PAR model was the only one studied that consistently provided theoretical Poisson's ratios in UCC that were positive, in agreement with several studies [10,17,19,38,39]. Furthermore, the anisotropic and asymmetric predictions of Poisson's ratios were similar to those measured in several studies; for example, the UCC Poisson's ratio  $\nu_{-13}$  was  $\sim 3\times$  greater than  $\nu_{-12}$  as compared to experimental values of  $5\text{--}6\times$  [10] and  $2\times$  [19], and the UT Poisson's ratio  $\nu_{+13}$  was  $\sim 2\times$  greater than  $\nu_{+12}$  as compared to an experimental value of  $2\times$  [13]. Although the UCC Poisson's ratios predicted by the 8-PAR model were substantially lower than the assumed values at 20% strain (Table 2), they were the same order of magnitude as those on the lower end of the reported values [10,19,38,39]. Thus, a limitation of the present study is the uncertainty in these predicted UCC Poisson's ratios, as the values assumed were based on experiments using a different tissues source than that assumed for the tensile properties.

The 8-PAR-B model, which did not explicitly include the traction-free boundary conditions in the nonlinear regression analysis, yielded different material constants which resulted in a poor theoretical prediction of UT in the 2 direction. Interestingly, the 8-PAR-B model's theoretical solution for UT stress in the 2 direction based on the assumed Poisson's ratios, as used in the nonlinear regression analysis, was nearly indistinguishable from the experimental curve (result not shown). It appears that the relatively large predicted value of  $\delta$  for the 8-PAR-B model, being an order of magnitude greater than the value for the 8-PAR model, appears to magnify the difference between the UT stress calculated using the assumed Poisson's ratios (as used in the nonlinear regression) and using the theoretical Poisson's ratios (as shown in Fig. 5). These results highlight both the importance of including the traction-free boundary condition equations in the regression analysis and checking the complete theoretical solution after the regression analysis is performed.

The approach adopted here is based upon a phenomenological model; a greater understanding of the structure–function relationship for articular cartilage could provide additional insight into degenerative processes and repair strategies. Towards this broader aim, there is insufficient experimental data to completely charac-



**Fig. 5 Prediction of the eight-parameter model that does not include the traction-free boundary condition equations (8-PAR-B) for the uniaxial tension (UT) response in the 2 direction. The predictions for the UT response in the 1 and 3 directions and the UCC response in the 1 direction are similar to those of the 8-PAR model shown in Fig. 3.**

terize the relationship between the material constants proposed in this study and features of the tissue's microstructure. However, since the present phenomenological model has been derived in accordance with Spencer's theory of fiber-reinforced anisotropy, it does allow us to rationally discuss the possible microstructural interpretations of the material constants. Here, several possible links between the 8-PAR model and microstructural features are presented.

First, the material constant  $\mu$  representing the isotropic matrix may be related to the proteoglycan component of the articular cartilage solid matrix; this material constant is a primary determinant of the tissue's compressive stiffness.

Second, the material constants ( $\gamma_1, \gamma_2, \gamma_3$ ) associated with the primary fibers may be related to the strength and three-dimensional (3D) distribution of collagen fibers. Interestingly, even though the assumed experimental data was proposed to be valid for the superficial region where it is thought that the collagen fibers lie primarily in the split-line direction, the preliminary analysis that neglected the fiber strength in the direction perpendicular to the articular surface (i.e.,  $\gamma_3$ ) did not produce reasonable predictions. This result may suggest a limitation of the present study, as the tensile properties in this direction are not well documented [14].

Third, the material constants ( $\delta, \phi_{12}, \phi_{13}, \phi_{23}$ ) associated with the secondary fibers may be related to more than one mechanism. As with the primary fiber strengths ( $\gamma_2, \gamma_3$ ), they may represent fibers that do not lie in the direction of the split-line direction reflecting a 3D distribution of fibers. Also, they may represent crosslinking mechanisms that affect tensile properties, such as collagen-specific crosslinks [40], small proteoglycans [41–43], or bridging fibrils [44]. For example, the latter study [44] proposed an architectural model consisting of cartilage leaves, composed of a fine structure of fibers, that bend towards and become parallel with the surface layer (similar to the classical Benninghoff model [45]). Interestingly, in the surface region those authors state "linking fibrils appeared to bind adjacent collagen leaves to one another"<sup>11</sup>; that observation may provide one microstructural interpretation of fibers (both primary and secondary) other than those aligned with the split-line direction.

Despite the lack of a definitive microstructural interpretation of the material constants used here, the results do suggest that bimodular strong interaction terms for the orthotropic scalar invariants facilitate modeling the asymmetry in Poisson's ratios that have been experimentally measured. From a phenomenological perspective, the secondary fibers in the present model are responsible

<sup>11</sup>Quote taken from p. 798 of Ref. [44].

for causing the large lateral contractions in UT that produce Poisson's ratios on the order of 1.0–2.0. In particular, in the UT solution in the 1 direction, the primary fibers in the 2 and 3 directions go into compression; however, the secondary fibers in the 1-2 and 1-3 planes go into tension. Thus, these secondary fibers in the 1-2 and 1-3 planes become mechanically active in a manner that allows them to aid in “contracting” the specimen in the 2 and 3 directions, respectively.

It is possible that other continuum and/or microstructural approaches may yield similar predictions of the anisotropic and asymmetric mechanical response of articular cartilage as adopted here. These approaches may include polyconvex models employing the strong interactions terms proposed in Ref. [34] or used in Ref. [27,36] for orthotropic materials, or anisotropic models based on a 3D distribution of fibers that extend the 2D approach of Ref. [46] such as the recent study of Ref. [47]. However, this study does appear to be the first to develop a polyconvex model that can provide a reasonable prediction for the stresses and Poisson's ratios measured for articular cartilage in UT and UCC in multiple directions.

A secondary aim of this paper was to obtain a relatively simple model using a minimum number of parameters needed to model the desired response. Using a minimum number of parameters facilitates parameter estimation from experimental datasets. In the models presented here, parameter estimates were insensitive to initial values used, as required by the nonlinear regression analysis. As mentioned earlier, many models were studied in addition to those presented in the “Results” section. For example, a ten-parameter model was based on the bimodular strain energy function Eq. (20) with different strengths (i.e.,  $\delta_{\pm 12} \neq \delta_{\pm 13} \neq \delta_{\pm 23}$ ) for the strong interaction terms. Although that model resulted in a lower error (0.721) than the 8-PAR model, the asymptotic correlation matrix and confidence interval statistics suggested that the ten-parameter model was over-parameterized. That conclusion is consistent with the observation that the parameters estimated for that ten-parameter model did depend on initial values, due to over-parameterization. Additional models derived as reduced forms of the 8-PAR model were studied; however, none of these models were able to model the anisotropic and asymmetric Poisson's ratios. For example, when the primary fiber strength  $\gamma_3$  was neglected, the secondary fiber angles  $\phi_{12}$  and  $\phi_{13}$  converged to values close to 0 deg in the regression, producing predictions very similar to those of the 4-PAR model.

There are limitations related to the assumed experimental dataset. First, the data used did not correspond to a complete set of UT and UCC experiments for a specific source of articular cartilage (i.e., anatomic site, species, age, etc.). However, it did describe a highly anisotropic and asymmetric mechanical response typical of cartilage. Second, the experimental dataset used assumed an isotropic UCC response. In preliminary studies that assumed an anisotropic UCC response, many anisotropic polyconvex strain energy functions generalizing  $W_o$  were used based on Ref. [36]. Although that formulation can model substantial UCC anisotropy, it introduced additional material constants that rendered the model over-parameterized given the assumed experimental data. Furthermore, there is little data with respect to the anisotropic UCC properties at large deformations. In a recent study [25], we have proposed that a model employing a bimodular anisotropic collagen network (using methods developed here) that restrains an isotropic proteoglycan matrix can explain several anisotropic UCC features observed in young bovine cartilage [10,17,19]. In order to address these limitations, we have recently developed experimental protocols to measure confined compression, unconfined compression, and torsion properties in an anisotropic manner at large deformations [19].

It is important to note that the desired accuracy of the stress constitutive equation depends on the application that it is being used for. For example, there is uncertainty regarding the accuracy needed to predict areas of peak stresses and, consequently, regions

of failure for articular cartilage of in vivo joints. One study that modeled joint contact using an idealized geometry found that a transversely isotropic model, as compared to an isotropic one, better predicts locations of peak stress that agree with injury location following impact [11]. In contrast, another study using experimental contact pressure measurements and finite-element analysis simulation using computed tomography-generated mesh geometry [48] found that predicted areas of maximum contact pressure may be more sensitive to surface mesh topology.

The primary motivation for the present work was to obtain accurate stress constitutive equations that are needed to conduct robust validation tests of the cartilage growth mixture models that model proteoglycan and collagen growth at different rates [49–54]. If these growth models can be validated for specific in vitro protocols, then it may be possible to accurately predict 3D geometry changes of graft tissue or tissue engineered constructs that are needed to repair a specific site. Indeed, difficulties associated with one current clinical repair strategy, osteochondral graft implantation, include the construction of a smooth convex joint surface and mismatch between donor and repair site thickness [55]. Current efforts are aimed at describing 3D geometry changes for explants grown in vitro. In order to accurately model, or predict, how thickness and diameter of a cylindrical implant would change during growth, one needs accurate anisotropic stress constitutive equations for tension and compression states. Indeed, current (unpublished) studies are finding substantially different predictions of 3D geometry of constructs grown in vitro using cartilage growth models employing different stress constitutive equations; the results depend on the degree of anisotropy assumed in the constitutive model.

## Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 0245709, the Office of Naval Research (Department of Navy), and the Donald E. Bently Center for Engineering Innovation.

## Appendix: Proof of Continuity and Stability Conditions

Here, it is shown that the strain energy function  $W_{BIM}$  defined in Eq. (20) satisfies the bimodular stress continuity and polyconvexity conditions provided that the restrictions in Eq. (21) are met. It suffices to prove these conditions in a general case; any term in Eq. (20) and its associated surface of discontinuity can be stated in general form as

$$W_a = \frac{1}{6} \gamma_a (\lambda_a^2 - 1)^3 = \frac{1}{6} \gamma_a (\mathbf{C} \cdot \mathbf{A} - 1)^3, \quad g_a = \lambda_a^2 - 1 = \mathbf{C} \cdot \mathbf{A} - 1 \quad (A1)$$

where  $\mathbf{A} = \mathbf{a} \otimes \mathbf{a}$  is a structural tensor defined by a fiber direction  $\mathbf{a}$ . It is convenient to express Eq. (A1) using indicial notation

$$W_a = \frac{1}{6} \gamma_a (C_{MN} A_{MN} - 1)^3, \quad g_a = C_{MN} A_{MN} - 1 \quad (A2)$$

**Bimodular Stress Continuity Condition.** The terms in the stress and elasticity tensors derived from Eq. (A2) are calculated as  $S_{AB} = 2\partial W / \partial C_{AB}$  and  $\mathbf{C}_{ABCD} = \partial S_{AB} / \partial C_{CD}$ , respectively, and are highlighted as follows

$$S_{AB} = \dots + \gamma_a (C_{MN} A_{MN} - 1)^2 A_{AB} + \dots$$

$$\mathbf{C}_{ABCD} = \dots + 2\gamma_a (C_{MN} A_{MN} - 1) A_{AB} A_{CD} + \dots \quad (A3)$$

Considering the surface of discontinuity defined by  $C_{MN} A_{MN} - 1 = 0$ , it is seen that  $S_{AB} = 0$  and  $\mathbf{C}_{ABCD} = 0$  on this surface. Thus, both the stress and elasticity tensors are continuous at the surface; i.e.,  $\mathbf{S} = \mathbf{S}_+ = \mathbf{S}_-$  and  $[[\mathbf{C}]] = \mathbf{C}_+ - \mathbf{C}_- = 0$ . Consequently, the bimodular stress continuity conditions Eq. (7) are satisfied with the choice of  $s(\mathbf{C}) = 0$ .

**Polyconvexity Condition.** Considering the sufficient conditions for polyconvexity stated in Eq. (16), it suffices to show that

the strain energy term  $W_a$  defined in Eqs. (A1) and (A2) is a convex function of  $\mathbf{F}$ . In indicial notation, the convexity condition requires

$$\frac{\partial^2 W_a}{\partial F_{iA} \partial F_{jB}} H_{iA} H_{jB} > 0 \quad \text{for all } H_{kC} \neq F_{kC}, \quad H_{kC} \neq 0 \quad (\text{A4})$$

where  $W_a$  can be written in terms of  $\mathbf{F}$  as

$$W_a = \frac{1}{6} \gamma_a (F_{kM} F_{kN} A_{MN} - 1)^3 \quad (\text{A5})$$

A straightforward differentiation in indicial notation leads to

$$\begin{aligned} \frac{\partial^2 W_a}{\partial F_{iA} \partial F_{jB}} &= \gamma_a (F_{kM} F_{kN} A_{MN} - 1)^2 \delta_{ij} A_{AB} \\ &+ 4 \gamma_a (F_{kM} F_{kN} A_{MN} - 1) F_{iP} A_{AP} F_{jQ} A_{QB} \quad (\text{A6}) \end{aligned}$$

Consequently, one obtains

$$\begin{aligned} \frac{\partial^2 W_a}{\partial F_{iA} \partial F_{jB}} H_{iA} H_{jB} &= \gamma_a (F_{kM} F_{kN} A_{MN} - 1)^2 H_{iA} A_{AB} H_{iB} \\ &+ 4 \gamma_a (F_{kM} F_{kN} A_{MN} - 1) F_{iP} A_{AP} F_{jQ} A_{QB} H_{iA} H_{jB} \quad (\text{A7}) \end{aligned}$$

This result can be expressed in direct notation in terms of the fiber direction  $\mathbf{a}$  defined by  $\mathbf{A} = \mathbf{a} \otimes \mathbf{a}$  and the fiber stretch  $\lambda_a$  as

$$\begin{aligned} \frac{\partial^2 W_a}{\partial \mathbf{F} \partial \mathbf{F}} \mathbf{H} \cdot \mathbf{H} &= \gamma_a (\lambda_a^2 - 1)^2 (\mathbf{H} \mathbf{a}) \cdot (\mathbf{H} \mathbf{a}) \\ &+ 4 \gamma_a (\lambda_a^2 - 1) [(\mathbf{F} \mathbf{a}) \cdot \mathbf{H}] \cdot [(\mathbf{F} \mathbf{a}) \cdot \mathbf{H}] \quad (\text{A8}) \end{aligned}$$

Since the scalar products  $(\mathbf{H} \mathbf{a}) \cdot (\mathbf{H} \mathbf{a})$  and  $[(\mathbf{F} \mathbf{a}) \cdot \mathbf{H}] \cdot [(\mathbf{F} \mathbf{a}) \cdot \mathbf{H}]$  are always positive, a necessary and sufficient condition for Eq. (A8) to be positive is that  $\gamma_a$  be positive when  $\lambda_a > 1$  and equal to zero when  $\lambda_a < 1$ , as defined in Eq. (21).

In conclusion, the bimodular stress continuity condition allows the material constant  $\gamma_a$  to jump across the surface of discontinuity, while the polyconvexity condition further requires that  $\gamma_a$  be a positive constant on the tensile side and equal to zero on the compressive side.

## References

- [1] Lai, W. M., Hou, J. S., and Mow, V. C., 1991, "A Triphasic Theory for the Swelling and Deformation Behaviors of Articular Cartilage," *J. Biomech. Eng.*, **113**, pp. 245–258.
- [2] Basser, P. J., Schneiderman, R., Bank, R. A., Wachtel, E., and Maroudas, A., 1998, "Mechanical Properties of the Collagen Network in Human Articular Cartilage as Measured by Osmotic Stress Technique," *Arch. Biochem. Biophys.*, **351**, pp. 207–219.
- [3] Venn, M. F., and Maroudas, A., 1977, "Chemical Composition and Swelling of Normal and Osteoarthritic Femoral Head Cartilage. I. Chemical Composition," *Ann. Rheum. Dis.*, **36**, pp. 121–129.
- [4] Mow, V. C., and Ratcliffe, A., 1997, "Structure and Function of Articular Cartilage and Meniscus," *Basic Orthopaedic Biomechanics*, V. C. Mow and W. C. Hayes, eds., Raven Press, New York, pp. 113–178.
- [5] Woo, S. L.-Y., Akeson, W. H., and Jemmot, G. F., 1976, "Measurements of Nonhomogeneous Directional Mechanical Properties of Articular Cartilage in Tension," *J. Biomech.*, **9**, pp. 785–791.
- [6] Woo, S. L.-Y., Lubock, P., Gomez, M. A., Jemmot, G. F., Kuei, S. C., and Akeson, W. H., 1979, "Large Deformation Nonhomogeneous and Directional Properties of Articular Cartilage in Uniaxial Tension," *J. Biomech.*, **12**, pp. 437–446.
- [7] Akizuki, S., Mow, V. C., Muller, F., Pita, J. C., Howell, D. S., and Manicourt, D. H., 1986, "Tensile Properties of Human Knee Joint Cartilage: I. Influence of Ionic Conditions, Weight Bearing, and Fibrillation on the Tensile Modulus," *J. Orthop. Res.*, **4**, pp. 379–392.
- [8] Soltz, M. A., and Ateshian, G. A., 2000, "A Conewise Linear Elasticity Mixture Model for the Analysis of Tension-Compression Nonlinearity in Articular Cartilage," *J. Biomech. Eng.*, **122**, pp. 576–586.
- [9] Laasanen, M., Toyras, J., Korhonen, R., Rieppo, J., Saarakkala, S., Nieminen, M., Hirvonen, J., and Jurvelin, J. S., 2003, "Biomechanical Properties of Knee Articular Cartilage," *Biorheology*, **40**, pp. 133–140.
- [10] Wang, C. C., Chahine, N. O., Hung, C. T., and Ateshian, G. A., 2003, "Optical Determination of Anisotropic Material Properties of Bovine Articular Cartilage in Compression," *J. Biomech.*, **36**(3), pp. 339–353.
- [11] Donzelli, P. S., Spilker, R. L., Ateshian, G. A., and Mow, V. C., 1999, "Contact Analysis of Biphasic Transversely Isotropic Cartilage Layers and Correlations with Tissue Failure," *J. Biomech.*, **32**(10), pp. 1037–1047.
- [12] Krishnan, R., Park, S., Eckstein, F., and Ateshian, G. A., 2003, "Inhomogeneous Cartilage Properties Enhance Superficial Interstitial Fluid Support and Frictional Properties, but do not Provide a Homogeneous State of Stress," *J. Biomech. Eng.*, **125**(5), pp. 569–577.
- [13] Chang, D. G., Lottman, L. M., Chen, A. C., Schinagl, R. M., Albrecht, D. R., Pedowitz, R. A., Brossman, J., Frank, L. R., and Sah, R. L., 1999, "The Depth-Dependent, Multi-axial Properties of Aged Human Patellar Cartilage in Tension," *Trans. Annu. Meet. - Orthop. Res. Soc.*, **24**, p. 644.
- [14] Korhonen, R. K., Toyras, J., Nieminen, M. T., Rieppo, J., Hirvonen, J., Helminen, H. J., and Jurvelin, J. S., 2001, "Effect of Ionic Environment on the Compression-Tension Nonlinearity of Articular Cartilage in the Direction Perpendicular to Articular Surface," *Trans. Annu. Meet. - Orthop. Res. Soc.*, **26**, p. 439.
- [15] Elliot, D. M., Narmoneva, D. A., and Setton, L. A., 2002, "Direct Measurement of the Poisson's Ratio of Human Patella Cartilage in Tension," *J. Biomech. Eng.*, **124**, pp. 223–228.
- [16] Charlebois, M., McKee, M. D., and Buschmann, M. D., 2004, "Nonlinear Tensile Properties of Bovine Articular Cartilage and Their Variation with Age and Depth," *J. Biomech. Eng.*, **126**, pp. 129–137.
- [17] Chahine, N. O., Wang, C. C., Hung, C. T., and Ateshian, G. A., 2004, "Anisotropic Strain-Dependent Material Properties of Bovine Articular Cartilage in the Transitional Range From Tension to Compression," *J. Biomech.*, **37**, pp. 1251–1261.
- [18] Huang, C. Y., Stankiewicz, A., Ateshian, G. A., and Mow, V. C., 2005, "Anisotropy, Inhomogeneity, and Tension-Compression Nonlinearity of Human Glenohumeral Cartilage in Finite Deformation," *J. Biomech.*, **38**(4), pp. 799–809.
- [19] Ficklin, T., Thomas, G., Chen, A., Sah, R., Davol, A., and Klisch, S., 2006, "Development of an Experimental Protocol to Measure Anisotropic Material Properties of Bovine Articular Cartilage," *Proceedings Summer Bioengineering Conference*, ASME, Amelia Island, FL.
- [20] Curnier, A., He, Q. C., and Zysset, P., 1995, "Conewise Linear Elastic Materials," *J. Elast.*, **37**, pp. 1–38.
- [21] Holzapfel, G. A., Gasser, T. C., and Ogden, R. W., 2004, "Comparison of a Multi-layer Structural Model for Arterial Walls with a Fung-Type Model, and Issues of Material Stability," *J. Biomech. Eng.*, **126**(2), pp. 264–275.
- [22] Baer, A. E., Laursen, T. A., Guilak, F., and Setton, L. A., 2004, "The Micro-mechanical Environment of Intervertebral Disc Cells Determined by a Finite Deformation, Anisotropic, and Biphasic Finite Element Model," *J. Biomech. Eng.*, **125**, pp. 1–11.
- [23] Klisch, S. M., 2006, "A Bimodular Theory for Finite Deformations: Comparison of Orthotropic Second-order and Exponential Stress Constitutive Equations for Articular Cartilage," *Biomech. Model. Mechanobiol.*, **5**(2–3), pp. 90–101.
- [24] Ball, J. M., 1977, "Convexity Conditions and Existence Theorems in Non-linear Elasticity," *Arch. Ration. Mech. Anal.*, **63**, pp. 337–403.
- [25] Klisch, S. M., Sah, R. L., and Davol, A., 2006, "Bimodular-Orthotropic-Polyconvex Strain Energy Functions for the Collagen-Proteoglycan Solid Matrix of Articular Cartilage," *Proceedings Summer Bioengineering Conference*, ASME, Amelia Island, FL.
- [26] Schroder, J., and Neff, P., 2003, "Invariant Formulation of Hyperelastic Transverse Isotropy Based on Polyconvex Free Energy Functions," *Int. J. Solids Struct.*, **40**, pp. 401–445.
- [27] Itskov, M., Ehret, A. E., and Mavrilas, D., 2006, "A Polyconvex Anisotropic Strain-Energy Function for Soft Collagenous Tissues," *Biomech. Model. Mechanobiol.*, **5**(1), pp. 17–26.
- [28] Wagner, D. R., Reiser, K. M., and Lotz, J. C., 2006, "Glycation Increases Human Annulus Fibrosus Stiffness in Both Experimental Measurements and Theoretical Predictions," *J. Biomech.*, **39**, pp. 1021–1029.
- [29] Spencer, A. J. M., 1984, *Continuum Theory of the Mechanics of Fiber-Reinforced Composites*, Springer, New York.
- [30] Klisch, S. M., and Lotz, J. C., 1999, "Application of a Fiber-Reinforced Continuum Theory to Multiple Deformations of the Annulus Fibrosus," *J. Biomech.*, **32**(10), pp. 1027–1036.
- [31] Elliott, D. M., and Setton, L. A., 2000, "A Linear Material Model for Fiber-Induced Anisotropy of the Annulus Fibrosus," *J. Biomech. Eng.*, **122**, pp. 173–179.
- [32] Eberlein, R., Holzapfel, G. A., and Schulze-Bauer, C. A., 2001, "An Anisotropic Constitutive Model for Annulus Tissue and Enhanced Finite Element Analyses of Intact Lumbar Disc Bodies," *Comput. Methods Biomech. Biomed. Eng.*, **4**, pp. 209–230.
- [33] Spencer, A. J. M., 1971, "Theory of Invariants," *Continuum Physics*, A. C. Eringen, ed., Academic, New York, pp. 240–353.
- [34] Schroder, J., Neff, P., and Balzani, D., 2005, "A Variational Approach for Materially Stable Anisotropic Hyperelasticity," *Int. J. Solids Struct.*, **42**, pp. 4352–4371.
- [35] Wagner, D. R., 2002, "A Mechanistic Strain Energy Function and Experimental Results for the Human Annulus Fibrosus," Ph.D. thesis, U. C. Berkeley, Berkeley, CA.
- [36] Itskov, M., and Aksel, N., 2004, "A Class of Orthotropic and Transversely Isotropic Hyperelastic Constitutive Models Based on a Polyconvex Strain Energy Function," *Int. J. Solids Struct.*, **41**, pp. 3833–3848.
- [37] Chang, D. G., 1999, "Structure and Function Relationships of Articular Cartilage in Osteoarthritis," Ph.D. thesis, University of California, San Diego, La Jolla, CA.
- [38] Jurvelin, J. S., Buschmann, M. D., and Hunziker, E. B., 1997, "Optical and

- Mechanical Determination of Poisson's Ratio of Adult Bovine Humeral Articular Cartilage," *J. Biomech.*, **30**, pp. 235–241.
- [39] Wong, M., Ponticciello, M., Kovanen, V., and Jurvelin, J. S., 2000, "Volumetric Changes of Articular Cartilage During Stress Relaxation in Unconfined Compression," *J. Biomech.*, **33**(9), pp. 1049–1054.
- [40] Williamson, A. K., Chen, A. C., Masuda, K., Thonar, E. J.-M. A., and Sah, R. L., 2003, "Tensile Mechanical Properties of Bovine Articular Cartilage: Variations with Growth and Relationships to Collagen Network Components," *J. Orthop. Res.*, **21**, pp. 872–880.
- [41] Broom, N. D., and Poole, C. A., 1983, "Articular Cartilage Collagen and Proteoglycans. Their Functional Interdependency," *Arthritis Rheum.*, **26**(9), pp. 1111–1119.
- [42] Vynios, D. H., Papageorgakopoulou, N., Sazakli, H., and Tsiganos, C. P., 2001, "The Interactions of Cartilage Proteoglycans with Collagens are Determined by Their Structures," *Biochimie*, **83**, pp. 899–906.
- [43] Hendrickx, S., Thomas, P., Preston, B., and Stanton, P., 2001, "Partial Characterization of Matrix Components Interacting with Cartilage Proteoglycans," *Arch. Biochem. Biophys.*, **390**(2), pp. 186–194.
- [44] Jeffery, A. K., Blunn, G. W., Archer, C. W., and Bentley, G., 1991, "Three-Dimensional Collagen Architecture in Bovine Articular Cartilage," *J. Bone Jt. Surg., Br. Vol.* **73-B**, 795–801.
- [45] Benninghoff, A., 1925, "Form und Bau der Gelenkknorpel in Ihren Beziehungen zur Funktion. Zweiter Teil: der Aufbau des Gelenkknorpels in Seinen Beziehungen zur Funktion," *Z. Zellforsch Mikrosk Anat.*, **2**, pp. 783–862.
- [46] Lanir, Y., 1983, "Constitutive Equations for Fibrous Connective Tissues," *J. Biomech.*, **16**, pp. 1–12.
- [47] Gasser, T. C., Ogden, R. W., and Holzapfel, G. A., 2006, "Hyperelastic Modelling of Arterial Layers with Distributed Collagen Fibre Orientations," *J. R. Soc., Interface*, **3**, pp. 15–35.
- [48] Anderson, A. F., Ellis, B. J., Maas, S. A., Peters, C. L., Ateshian, G. A., and Weiss, J. A., 2006, "Experimental Measurement and Finite Element Prediction of Cartilage Contact Pressures in the Human Hip," *Proceedings ASME Summer Bioengineering Conference*, Amelia Island, FL.
- [49] Klisch, S. M., Sah, R. L., and Hoger, A., 2000, "A Growth Mixture Theory for Cartilage," *Mechanics in Biology*, J. Casey and G. Bao, eds., ASME, AMD-242, New York, pp. 229–242.
- [50] Klisch, S. M., Van Dyke, T., and Hoger, A., 2001, "A Theory of Volumetric Growth for Compressible Elastic Materials," *Math. Mech. Solids*, **6**, pp. 551–575.
- [51] Klisch, S. M., and Hoger, A., 2003, "Volumetric Growth of Thermoelastic Materials and Mixtures," *Math. Mech. Solids*, **8**, pp. 377–402.
- [52] Klisch, S. M., Chen, S. S., Sah, R. L., and Hoger, A., 2003, "A Growth Mixture Theory for Cartilage with Applications to Growth-Related Experiments on Cartilage Explants," *J. Biomech. Eng.*, **125**, pp. 169–179.
- [53] Klisch, S. M., Sah, R. L., and Hoger, A., 2005, "A Cartilage Growth Mixture Model for Infinitesimal Strains: Solutions of Boundary-Value Problems Related to In Vitro Growth Experiments," *Biomech. Model. Mechanobiol.*, **3**(4), pp. 209–223.
- [54] Davoli, A., Bingham, M. S., Sah, R. L., and Klisch, S. M., 2007, "A Nonlinear Finite Element Model of Cartilage Growth," *Biomech. Model. Mechanobiol.*, in press.
- [55] Smith, G. D., Knutsen, G., and Richardson, J. B., 2005, "A Clinical Review of Cartilage Repair Techniques," *J. Bone Jt. Surg., Am. Vol.* **87B**(4), pp. 445–449.

**SBC2007-175327**

## EXTENDED TWO COMPARTMENTAL SWELLING STRESS MODEL AND ISOTROPIC CAUCHY STRESS EQUATION FOR ARTICULAR CARTILAGE PROTEOGLYCANS

Sevan R. Oungouljian (1), Silvia S. Chen (2), Andrew Davol (1), Robert L. Sah (2),  
Stephen M. Klisch (1)

(1) Mechanical Engineering Department  
California Polytechnic State University  
San Luis Obispo, CA

(2) Department of Bioengineering  
University of California, San Diego  
La Jolla, CA

### INTRODUCTION

Proteoglycans (PGs), a constituent of cartilaginous tissues, have a negative fixed charge (FC) that causes an intratissue swelling stress [1]. This swelling stress is thought to balance tensile stress in the collagen network and contribute to the aggregate modulus of articular cartilage (AC) [1]. Stress constitutive equations that accurately characterize mechanical behavior of individual tissue constituents are crucial for the development of accurate total tissue models. The goal of this study is to extend the range of an existing two compartmental model for PG swelling stress by Basser et al. [1], and develop a continuum level equation for PG Cauchy stress. Specifically, the first aim is to increase the accuracy of the two compartmental model proposed in [1], to a lower range of FC density (FCD) typically found in bovine calf AC. The second aim is to use the extended model to develop a continuum level strain energy function and associated isotropic PG Cauchy stress constitutive equation.

### METHODS

The two compartmental swelling stress model [1] is given as

$$FCD_{eff} = FCD_{tot} / m_{ef}^w, \quad (1)$$

$$m_{ef}^w = m^{ww} - m^{dw} - m_{if}^w, \quad (2)$$

$$m_{if}^w = (0.726 + 0.538 \exp(-0.258\pi^{PG})), \quad (3)$$

$$\pi^{PG} = 3.51FCD_{eff} + 19.30FCD_{eff}^2, \quad (4)$$

where  $m_{ef}^w$  and  $m_{if}^w$  are extrafibrillar and intrafibrillar water masses (mg) respectively,  $m^{ww}$  and  $m^{dw}$  are tissue wet and dry masses (mg) respectively,  $FCD_{eff}$  and  $FCD_{tot}$  are effective (normalized to  $m_{ef}^w$ ) and total (normalized to total water mass) FCD (mEq/g)

respectively, and  $\pi^{PG}$  is PG swelling stress (atm).

Because bovine calf AC explants typically have a low FCD, additional data for the relation between  $\pi^{PG}$  and  $FCD_{tot}$  were used [2]. Data from both [1] and [2] were fit to the exponential relationship

$$IF \ FCD_{eff} \leq g: \quad \pi^{PG} = -a(1 - \exp(b FCD_{eff})), \quad (5a)$$

$$IF \ FCD_{eff} \geq g: \quad \pi^{PG} = c \exp(d FCD_{eff}) + f, \quad (5b)$$

using nonlinear regression (MATLAB), where  $a - g$  are material constants.

Experimental constituent mass normalized by tissue wet weight (m/WW) data for AC harvested from the patellofemoral groove of bovine calves were taken from a previous study (Table 1) [3]. Specimens were sliced ( $0.25\text{mm} < \text{thickness} < 0.5\text{mm}$ ) into three successive layers: superficial, first middle, and second middle zone layers [3].

**Table 1. Constituent mass normalized to tissue wet weight (m/WW) for bovine calf AC explants at three depths: superficial, first middle, and second middle zone layers [3]. All values given as mean  $\pm$  one standard deviation.**

Experimental measurements of m/WW (mg/g)			
Layer	Water	Collagen	Proteoglycan
Superficial	875 $\pm$ 258	84 $\pm$ 16	29 $\pm$ 4
First Middle	866 $\pm$ 244	108 $\pm$ 16	35 $\pm$ 5
Second Middle	842 $\pm$ 126	134 $\pm$ 28	38 $\pm$ 4

Total FC ( $FC_{tot}$ ) was assumed to be proportional to glycosaminoglycan (GAG) mass ( $m^{GAG}$ ) [4]:

$$FC_{tot} = 3.76(458/502)m^{GAG}. \quad (6)$$

because  $m^{GAG}$  [3] was measured from dimethylene blue assay using chondroitin sulfate C (502 g/mol) as a standard, the molecular weight of dissociated CS disaccharide being 458 g/mol, the conversion factor (458/502) is included in Eq. (6).

Confined compression (CC) tests were simulated to evaluate  $\pi^{PG}$  by solving the extended two compartmental model Eqs. (1) – (3) & (5) numerically (MATLAB). Extended model values were calculated at 0%, 15%, 30%, and 45% strain levels. Constituent apparent densities (constituent mass per total tissue volume) were calculated from reference densities and prescribed strain.

Dependence of  $\pi^{PG}$  on each constituent apparent density was tested by varying m/WW for PG, water, and collagen constituents over a range of mean  $\pm 2$  standard deviations.  $\pi^{PG}$  was seen to have a greater dependence on PG and water, rather than collagen, apparent densities over the tested range. From these results, two polynomial continuum level isotropic PG Cauchy stress equations, the first dependent on PG apparent density Eq. (7a) and the second dependent on both PG and water apparent densities Eq. (7b), were selected for further analysis and are given as

$$\underline{T}^{PG} = \alpha_1 (\rho^{PG})^{\alpha_2} \underline{I}, \quad (7a)$$

$$\underline{T}^{PG} = \alpha_3 (\rho^{PG})^{\alpha_4} (\rho^w)^{\alpha_5} \underline{I}, \quad (7b)$$

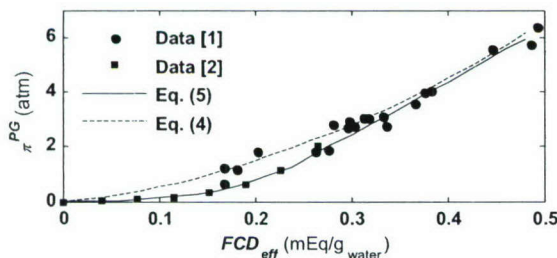
where  $\underline{T}^{PG}$  is PG Cauchy stress ( $\pi^{PG}$  normalized by total tissue area),  $\rho^{PG}$  is PG apparent density,  $\underline{I}$  is the identity tensor, and  $\alpha_1 - \alpha_5$  are material coefficients. Eqs. (7a) & (7b) were separately fit to the combined CC results from the extended two compartmental model for mean specimens in each of the surface, first middle, and second middle layers.

The swelling stress predicted by the commonly used ideal Donnan model was also computed at each specimen configuration, assuming the external solution ionic concentration remained constant at 0.15M NaCl [4]. Swelling stress in the tissue water area was normalized to total tissue area to obtain  $\underline{T}^{PG}$ .

Predictive capabilities of Eqs. (7a) & (7b) and the ideal Donnan model were assessed by computing  $R^2$  values from the comparison of each model to the extended two compartmental model [4].

## RESULTS

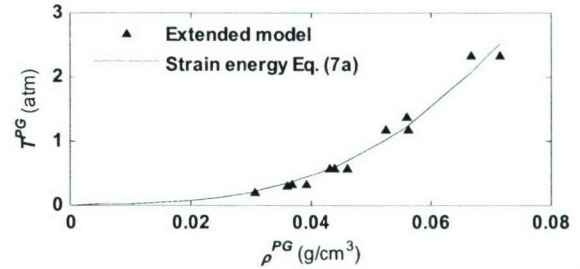
The material coefficients for the extended  $\pi^{PG}$  and  $FCD_{EF}$  relation, Eq. (5) were  $a = -0.0375$ ,  $b = 15.14$ ,  $c = 155.1$ ,  $d = 0.120$ ,  $f = -158.3$ , &  $g = 0.230$ , and resulted in  $R^2 = 0.96$  (Figure 1).



**Figure 1. Comparison of proteoglycan swelling stress ( $\pi^{PG}$ ) to effective fixed charge density ( $FCD_{eff}$ ) relation between extended Eq. (5) and Basser Eq. (4) [1] [2].**

The material coefficients for the strain energy based Cauchy stress equations, Eqs. (7a) & (7b), were  $\alpha_1 = 6020$ ,  $\alpha_2 = 2.95$ ,

$\alpha_3 = 4850$ ,  $\alpha_4 = 2.87$ , &  $\alpha_5 = 0.00$  and resulted in identical  $R^2$  values,  $R^2 = 0.97$ , when compared to the extended two compartmental model (Figure 2). The ideal Donnan model resulted in  $R^2 = 0.92$ .



**Figure 2. PG Cauchy stress ( $\underline{T}^{PG}$ ) from PG density ( $\rho^{PG}$ ) dependent strain energy based Eq. (7a) and extended two compartmental model Eqs. (1)-(3) & (5) during simulated CC of mean samples at three depths: superficial, first middle, and second middle zone layers [3]. Extended model values calculated at 0%, 15%, 30%, and 45% strain.**

## DISCUSSION

The results indicate that the two compartmental model swelling stress model was effectively extended to accurately describe PG swelling stress for low FCD levels without compromising overall accuracy. The results also suggest that the continuum level Eq. (7a) can provide a good estimate of PG Cauchy stress for specimens typically found in the surface and middle layers of bovine calf AC. Eq. (7a) only depends on PG apparent density, rather than the Jacobian of the PG deformation gradient tensor [5], therefore eliminating dependence on the tissue reference configuration and allowing application to AC specimens with different reference configuration densities. Additionally, Eq. (7a) is derived from a strain energy function that meets polyconvexity requirements when  $\alpha_1 \geq 0$  &  $\alpha_2 \geq 0$ , and is subsequently known to satisfy material stability criterion.

## ACKNOWLEDGEMENTS

Supported by NSF (AD, RS, SK), NIH (RS), the ONR (SK), and HHMI (RS).

## REFERENCES

1. Bassar, P. J., Schneidman, R., Bank, R. A., Wachtel, E., Maroudas, A., 1998, "Mechanical Properties of the Collagen Network in Human Articular Cartilage as Measured by Osmotic Stress Technique," *Archives of Biochemistry and Biophysics*, Vol. 351, No. 2, pp. 207-219.
2. Williams, R. P. W., Comper, W. D., 1990, "Osmotic flow caused by polyelectrolytes," *Biophys Chem*, 36, pp. 223-234.
3. Asanbaeva, A., McGowan, K. B., Masuda, K., Klisch, S. M., Thonar, E. J-MA., Sah, R. L., 2004, "Mechanisms of cartilage growth: Alteration of function and composition *in vitro* by deposition of collagen and proteoglycan matrix components," *Transactions of the ORS*.
4. Buschmann, M. D., Grodzinsky, A. J., 1995, "A Molecular Model of Proteoglycan-Associated Electrostatic Forces in Cartilage Mechanics," *J Biomech Eng*, 117, pp. 179-192.
5. Klisch, S. M., Sah, R. L., Davol, A., 2006, "Bimodular-orthotropic-polyconvex strain energy functions for the collagen-proteoglycan solid matrix of articular cartilage," *Transactions of the ASME, SBC*.

**SBC2007-175548**

## A CARTILAGE GROWTH MIXTURE MODEL WITH COLLAGEN REMODELING: VALIDATION PROTOCOLS

**Stephen M. Klisch (1), Anna Asanbaeva (2), Sevan R. Oungouljian (1), Eugene J-MA. Thonar (3,4), Koichi Masuda (3), Andrew Davol (1), Robert L. Sah (2)**

(1) Department of Mechanical Engineering  
California Polytechnic State University  
San Luis Obispo, CA

(2) Department of Bioengineering  
University of California, San Diego  
La Jolla, CA

(3) Departments of Biochemistry and Orthopedic  
Surgery  
Rush Medical Center  
Chicago, IL

(4) Department of Internal Medicine  
Rush Medical Center  
Chicago, IL

### INTRODUCTION

The attainment of specific design goals related to tissue structure and function is likely critical to the development of a consistently successful strategy for articular cartilage (AC) repair. The long-term goal of this research is to develop a cartilage growth mixture (CGM) model that may be used with experimental approaches to improve AC repair strategies. The CGM employs a "stress balance" hypothesis: at equilibrium, the crosslinked collagen (COL) network restrains the proteoglycan (PG) swelling pressure generated by fixed charge density. The specific objectives are to: 1) refine a previous CGM model [1] to include COL remodeling; 2) develop analytical protocols for model validation that are based on experimental data from select *in vitro* growth protocols; and 3) test CGM model hypotheses.

### METHODS

The superscripts p and c designate the PG and COL constituents, respectively. The CGM model employs a mixture of a water constituent and a growing solid matrix (SM) composed of two growing elastic constituents, PG and COL. The kinematics describe the evolution of a stress-free configuration of a homogeneous SM element; due to the homogeneity assumptions used here, the element corresponds to the entire tissue explant. Configurations  $\kappa_R$ ,  $\kappa_M$ , and  $\kappa_G$  specify unloaded stress-free configurations of the SM element before growth, after mass deposition, and after mass deposition and remodeling, respectively (Fig. 1). The model assumes an *immobility constraint*: all of the PG and COL molecules are bound to the SM, so that their total deformation gradient tensors  $\mathbf{F}^p$  and  $\mathbf{F}^c$  are equal to the SM deformation gradient tensor  $\mathbf{F}$ . During growth and remodeling of an unloaded element,  $\mathbf{F}$  is decomposed into tensors  $\mathbf{F}_m$  due to mass deposition and  $\mathbf{F}_r$  due to COL remodeling (Fig. 1).  $\mathbf{F}_m$  describes the evolution of the SM stress-free configuration due to mass deposition

and is decomposed into constituent elastic growth ( $\mathbf{F}_{eg}^p, \mathbf{F}_{eg}^c$ ) and mass growth ( $\mathbf{F}_g^p, \mathbf{F}_g^c$ ) tensors.  $\mathbf{F}_r$  describes the change in the SM stress-free configuration due to COL remodeling. Also, a superposed elastic deformation  $\mathbf{F}_l$  is applied in mechanical experiments. Using the immobility constraint one obtains:

$$\mathbf{F} = \mathbf{F}_l \mathbf{F}_r \mathbf{F}_m = \mathbf{F}_l \mathbf{F}_r \mathbf{F}_{eg}^p \mathbf{F}_g^p = \mathbf{F}_l \mathbf{F}_r \mathbf{F}_{eg}^c \mathbf{F}_g^c. \quad (1)$$

Consequently, the total *elastic tensors* of the PG and COL constituents ( $\mathbf{F}_e^p, \mathbf{F}_e^c$ ) are decomposed as

$$\mathbf{F}_e^p = \mathbf{F}_l \mathbf{F}_r \mathbf{F}_{eg}^p, \quad \mathbf{F}_e^c = \mathbf{F}_l \mathbf{F}_r \mathbf{F}_{eg}^c. \quad (2)$$

Here, the growth tensors are assumed to be isotropic:

$$\mathbf{F}_g^p = [J_g^p]^{1/3} \mathbf{I}, \quad \mathbf{F}_g^c = [J_g^c]^{1/3} \mathbf{I}, \quad (3)$$

where  $J_g^p = \det \mathbf{F}_g^p$  and  $J_g^c = \det \mathbf{F}_g^c$ . The balance of mass equations for the constituents are decomposed into elastic continuity and growth continuity equations. Here, the growth rate is assumed constant; consequently, the growth continuity equations allow the calculation of the mass growth tensors from mass measurements:

$$J_g^p = 1 + \Delta m / m_0^p, \quad J_g^c = 1 + \Delta m / m_0^c, \quad (4)$$

where  $\Delta m$  is the change in mass and  $m_0$  is the initial mass. Upon assuming that the SM is homogeneous and stress-free in  $\kappa_R$ ,  $\kappa_M$ , and  $\kappa_G$ , the equations of motion at equilibrium ( $\text{div} \mathbf{T}^s = \mathbf{0}$ ) reduce to

$$\mathbf{T}^s = \mathbf{T}^p + \mathbf{T}^c = \mathbf{0}. \quad (5)$$

PG and COL stress constitutive equations are defined, based on recent studies [2,3], relative to their respective stress-free reference configurations  $\kappa^p_0$  and  $\kappa^c_0$ . Here,  $\kappa^p_0$  coincides with  $\kappa_R$  while a COL elastic swelling strain  $\mathbf{F}^c_0$  is introduced that maps  $\kappa^c_0$  to  $\kappa_R$  (Fig. 1). Due to the assumption that the PG and COL stresses only depend on their respective elastic tensors, in addition to a COL stress dependence on a scalar remodeling factor  $\chi$ , general stress constitutive equations are defined with respect to  $\kappa_R$ :

$$\mathbf{T}^p = \hat{\mathbf{T}}^p_{\kappa_R}(\mathbf{F}^p_c), \quad \mathbf{T}^c = \hat{\mathbf{T}}^c_{\kappa_R}(\chi, \mathbf{F}^c_c \mathbf{F}^c_0). \quad (6)$$

Due to a lack of comprehensive data, each of the COL material constants is changed by the same multiplicative remodeling factor  $\chi$ .

Experimental measurements of wet weight, COL mass, glycosaminoglycan (GAG) mass, COL specific pyridinoline (PYR) crosslink content, and tensile modulus in a direction ~ perpendicular to the local split-lines were obtained from a previous study [4]. Bovine calf patellofemoral groove AC explants (n=9-21 from 5 animals) were harvested from three successive layers: (S) superficial ~0.4 mm thick; (M1) first middle zone ~0.25 mm thick; and (M2) second middle zone ~0.25 mm thick. Some blocks were analyzed immediately (control groups S-D0, M1-D0, M2-D0). Other blocks were incubated for 13 days in medium (DMEM supplemented with FBS and ascorbate) (experimental groups S-FBS, M1-FBS, M2-FBS) or medium with 0.1mM  $\beta$ -aminopropionitrile (BAPN) (experimental groups S-BAPN, M1-BAPN, M2-BAPN). The averaged values of the control and experimental groups are used to quantify growth.

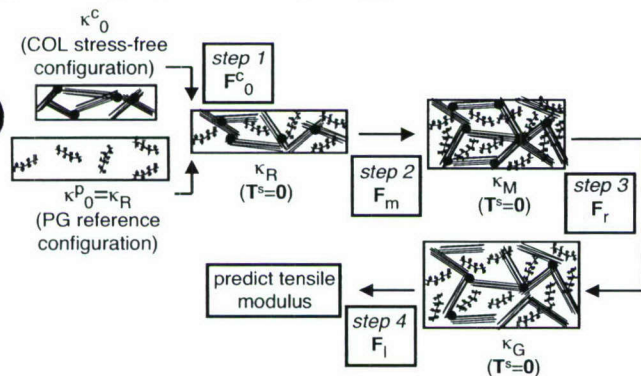


Figure 1. Analysis steps used to validate the CGM model.

Four analysis steps are used to quantify growth for each experimental group (Fig. 1). In step 1 the PG and COL material constants are specified for each control group and  $\mathbf{F}^c_0$  is calculated using (5-6). In step 2 the mass growth and elastic growth tensors are determined using experimental masses and (4-6). In step 3 the COL remodeling factor  $\chi$  (and consequently  $\mathbf{F}_r$ ) is determined so that the theoretical and experimental volumes agree for the final grown configuration  $\kappa_G$ . In step 4 validation tests are performed. The uniaxial tension boundary-value problem is solved and the secant Young's modulus at 20% strain is calculated for each experimental group. A paired t-test is performed between the sets of theoretical and experimental values of tensile modulus of the experimental groups. The relationship between  $\chi$  and PYR content is analyzed using linear regression with a t-test analysis of the regression slope.

## RESULTS

Theoretical values of tensile modulus predicted by the CGM model are within  $\pm 1$  standard deviations of experimental means; there

is not a significant difference between theoretical and experimental values ( $p=0.40$ ,  $n=6$ ) (Fig. 2-left). The theoretical remodeling factor  $\chi$  required to match final tissue volume exhibits a positive trend with PYR content that is not significant when pooling all experimental groups ( $R^2=0.37$ ,  $p=0.20$ ,  $n=6$ ). However, the trend exhibits a strong correlation when pooling the S and M1 experimental groups ( $R^2=1.00$ ,  $p<0.0001$ ,  $n=4$ , Fig. 2-right).

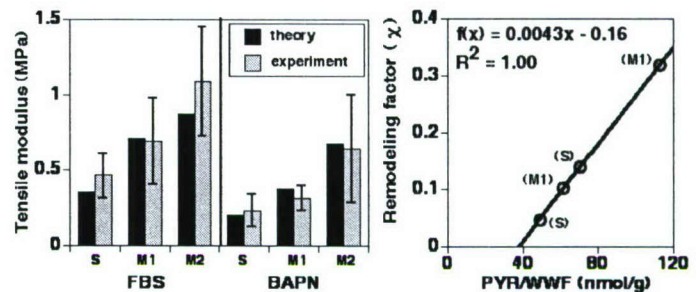


Figure 2. (Left) Theoretical and experimental (mean  $\pm 1$  S.D.) values of tensile modulus are not significantly different ( $p=0.40$ ,  $n=6$ ). (Right) The theoretical remodeling factor is correlated with PYR content when pooling the S and M1 groups ( $p<0.0001$ ,  $n=4$ ).

## DISCUSSION

Validation is obtained in three manners: (1) the CGM model is capable of matching final tissue composition and volume via specification of the mass growth tensors and remodeling factor  $\chi$ ; (2) upon determining  $\chi$  to match the grown tissue volume, the CGM model provides theoretical predictions of tensile modulus for three layers (S, M1, M2) and two *in vitro* growth protocols (FBS, BAPN) that do not significantly differ from experimental values; and (3)  $\chi$  is correlated with experimental measures of PYR crosslink content. The results of this study suggest that a CGM model may help to guide *in vitro* growth protocols for AC tissue via the *a priori* prediction of construct geometric and biomechanical properties.

## ACKNOWLEDGEMENTS

Supported by NSF (SK, AD, RS), ONR (SK), NIH (ET, KM, RS), HHMI (RS).

## REFERENCES

1. Klisch, S. M., Chen, S. S., Sah, R. L., and Hoger, A., 2003, "A growth mixture theory for cartilage with applications to growth-related experiments on cartilage explants," *J Biomech Eng*, 125, pp. 169-179.
2. Klisch, S. M., 2007, "A bimodular polyconvex anisotropic strain energy function for articular cartilage," *J Biomech Eng*, in press.
3. Oungoulian, S. R., Chen, S. S., Davol A., Sah, R. L., Klisch, S. M., 2007, "Extended two compartmental swelling stress model and isotropic Cauchy stress equation for articular cartilage proteoglycans," *ASME Summer Bioengineering Conference*, in review.
4. Asanbaeva, A., McGowan, K. B., Masuda, K., Klisch, S. M., Thonar, E. J.-M. A., and Sah, R. L., 2004, "Mechanisms of cartilage growth: alteration and function and composition *in vitro* by deposition of collagen and proteoglycan matrix components," *Trans Orthop Res Soc*, 29, p. 554.

**SBC2007-175965**

## **INVESTIGATION OF CARTILAGE BIOMECHANICAL PROPERTIES: DEPENDENCE ON STRAIN, DIRECTION, AND BIOCHEMICAL COMPOSITION**

**Gregory C. Thomas (1), Timothy P. Ficklin (1), James C. Barthel (1), Anna Asanbaeva (2), Eugene J-MA. Thonar (3,4), Koichi Masuda (3), Albert C. Chen (2), Robert L. Sah (2), Andrew Davol (1), Stephen M. Klisch (1)**

(1) Mechanical Engineering Department  
California Polytechnic State University  
San Luis Obispo, CA

(2) Department of Bioengineering  
University of California, San Diego  
La Jolla, CA

(3) Departments of Biochemistry and Orthopedic  
Surgery  
Rush Medical Center  
Chicago, IL

(4) Department of Internal Medicine  
Rush Medical Center  
Chicago, IL

### **INTRODUCTION**

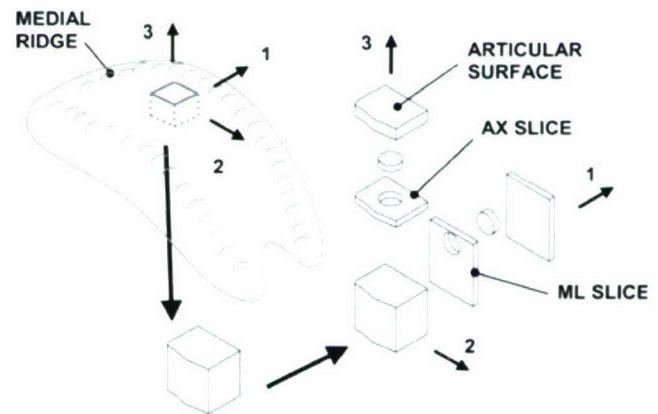
Articular cartilage (AC) serves as the major load bearing material within synovial joints and provides a low friction and wear resistant interface. As an avascular tissue, AC lacks the ability to repair structural damage or degeneration. Thus, the need for replacement tissue was a motivating factor in the development of cartilage tissue engineering. Recently, a finite element model (FEM) of cartilage growth [1] has been developed to simulate various growth conditions such as *in vitro* (outside the body) tissue growth experiments. In order to validate growth laws used in the FEM, empirical measurements of AC properties (mechanical and biochemical) before and after *in vitro* growth are needed. The goal of this study is to design protocols to comprehensively quantify the biomechanical structure-function relations of AC.

The specific aims of this project were to: (1) measure the direction and strain dependent biomechanical properties of AC before and after *in vitro* growth, and (2) quantify the relationships between mechanical properties and biochemical content.

### **METHODS**

#### **Harvesting**

Newborn (1-3 weeks old) bovine calf AC explants were harvested from the patellofemoral groove. For control groups, cylindrical disc test specimens were harvested from directions 1 and 3 as shown in Figure 1. Following a previously developed harvesting protocol [2], slices normal to the medial-lateral (ML) and axial (AX) directions were placed on a freezing stage and cut with a microtome to a thickness of 1 mm. A stainless steel punch was then used to extract 3.2 mm diameter discs at a mean depth of 2 mm below the articular surface

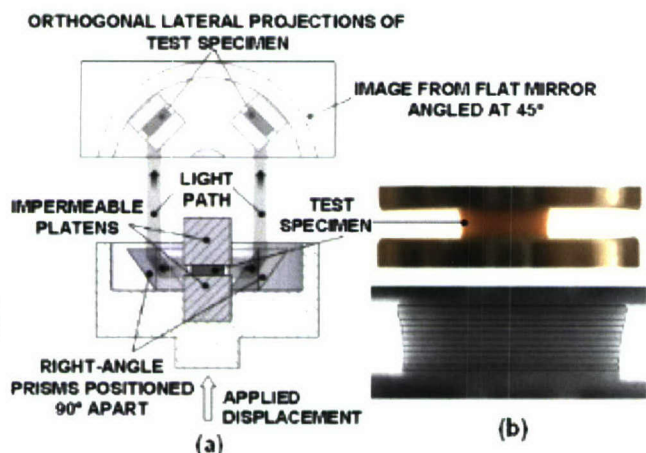


**Figure 1. Illustration of the harvesting procedure. For the control group, discs were harvested from two orthogonal directions.**

For an *in-vitro* group (AX-D14), cartilage slices from the AX direction were cut on a microtome and then with a scalpel into rectangular slices (10 x 6 x 0.75 mm) in order to retain the directionality of the specimen. After geometry and wet weight measurements, slices were then placed in petri dishes, submerged in a solution of FBS + ascorbate and incubated at 37°C for a period of 14 days in accord to existing cartilage growth protocols [3]. After 14 days, the slices were re-measured for geometry and wet weight, and then cartilage discs (4.8 mm diameter) were punched from each slice.

## Mechanical Tests – Unconfined Compression

Among a series of mechanical test protocols [2], both control and *in vitro* groups were subjected to an unconfined compression (UCC) test to measure Young's modulus and two orthogonal Poisson's ratios at 15, 30, and 45% compressive strains. In this experiment, a test specimen is submerged in phosphate buffered saline and compressed between two impermeable platens (Figure 2). A load cell (not shown) attached to the top platen measures the reaction force, which is used along with the axial displacement to calculate the equilibrium stress and secant Young's modulus ( $E$ ). In addition, two prisms are placed 90° apart and simultaneously reflect orthogonal views of the test sample. The reflected images are captured with a digital camera and processed in MATLAB to calculate Poisson's ratios. A MATLAB script was created to identify the test sample and determine the disc diameter by scanning through each pixel row and finding the location of greatest change in pixel intensity. With images of the sample at 0, 15, 30, and 45 % compressive strain, the diametral increase (transverse strain) and measured axial strain were used to calculate Poisson's ratios,  $v_{ij}$  ( $i$ =direction of axial strain,  $j$ =direction of transverse strain).



**Figure 2. Unconfined Compression Experiment. (a) Cross sectional illustration of the experiment. (b) An image taken during the experiment (top) is processed in MATLAB to measure the average disc diameter (bottom) with black lines indicating the measured diameter.**

## Biochemistry

After mechanical testing, each sample was measured for biochemical content. Using established protocols [3], samples were lyophilized, measured for dry weight, and then digested in a solution of Proteinase K. The digested solutions were then analyzed to quantify the content of collagen (COL), glycosaminoglycans (GAG), DNA, and collagen specific pyridinoline crosslinks (PYR).

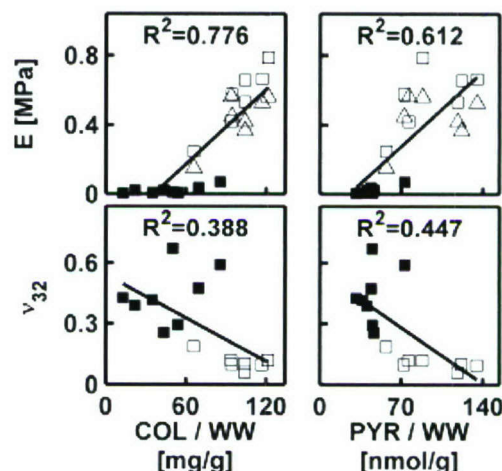
## RESULTS

From a statistical analysis, the measured mechanical properties were shown to have no significant dependence on strain. Furthermore, only Poisson's ratios exhibited anisotropic behavior (direction dependence). For example, at 30% strain,  $v_{13}$  ( $0.23 \pm 0.13$ ) (mean  $\pm$  s.d.) was significantly larger than  $v_{32}$  ( $0.12 \pm 0.04$ ), ( $p < 0.05$ ).

The biomechanical properties also showed significant changes after *in vitro* growth. After growth, the secant Young's modulus decreased and Poisson's ratios increased; indicating a weaker tissue sample. For example, at 30% strain the secant Young's modulus

decreased from  $0.50 \pm 0.19$  MPa to  $0.02 \pm 0.02$  MPa, and Poisson's ratio  $v_{31}$  increased from  $0.14 \pm 0.08$  to  $0.46 \pm 0.20$  after growth. In addition, all measured biochemistry content (COL, GAG, DNA, PYR) significantly decreased after growth ( $p < 0.05$ ).

ML, AX, and AX-D14 samples were then pooled together to test for structure-function relations between mechanical properties and biochemical content. From statistical analyses, Young's modulus and Poisson's ratios had the highest correlation to COL ( $p < 0.0001$ , 0.05 respectively) and PYR content ( $p < 0.01$ , 0.05) (Figure 3).



**Figure 3. Young's modulus ( $E$ ) and Poisson's ratio  $v_{32}$  at 30% strain were correlated with COL and PYR content (normalized to wet weight) for ML ( $\triangle$ ), AX ( $\square$ ), and AX-D14 ( $\blacksquare$ ) samples.**

## DISCUSSION

In order to gain insight on the structure-function relations of AC, mechanical test protocols incorporating novel methods of simultaneously measuring two orthogonal Poisson's ratios were developed. After measuring the biomechanical properties, statistical analyses were performed to investigate the dependence of mechanical properties on strain, direction, and biochemical content. As part of an on-going study, future work will include validating the cartilage growth FEM with empirical results from this and other *in vitro* growth experiments.

## ACKNOWLEDGEMENTS

Funding was received from NSF (RS, AD, SK), NIH (ET, KM, RS), and the Office of Naval Research (Department of Navy) (SK).

## REFERENCES

1. Davol A, Bingham, MS, Sah RL, Klisch SM. A nonlinear finite element model of cartilage growth under *in vitro* dynamic compression. *Biomechanics and Modeling in Mechanobiology*. 2007 (in press).
2. Ficklin TP, Thomas GC, Asanbaeva A, Chen AC, Sah RL, Davol A, Klisch SM. Development of an experimental protocol to measure anisotropic material properties of bovine articular cartilage. Transactions of the ASME Summer Bioengineering Conference, 2006.
3. Asanbaeva A, Masuda K, Thonar EJ, Klisch SM, Sah RL. Mechanisms of cartilage growth: modulation of balance between proteoglycan and collagen *in vitro* using chondroitinase ABC. *Arthritis and Rheumatism*, 56(1): 188-198, 2007.

**Delay and disruption tolerant network implementation with  
reconfigurable hardware support**

Project Investigator:

Albert A. Liddicoat  
Electrical Engineering and Computer Engineering  
California Polytechnic State University  
San Luis Obispo, California

**Introduction:** Today there are many applications that require high-performance reliable data networks that are robust against delay, network infrastructure, and node and link availability. These applications include military battle field networks, broad wireless sensor networks such as those used by the Navy for surveillance and scientific study, deep space exploration telemetry networks, and active RF-ID sensor networks. The fundamental internet protocols, Transmission Control Protocol (TCP) and Internet Protocol (IP), relies on several basic assumptions about its environment in order to function usefully; TCP/IP networks require short roundtrip network delays, low bit error rates, and guaranteed end-to-end connectivity: characteristics which most fixed terrestrial network environments provide. Today, the network of computers that make up the internet are connected together through low-delay, continuous end-to-end connections. TCP/IP will fail to function in more hazardous environments where networks have long link delays and temporary disruptions. The need to share information in less predictable networking environments, such as deep space networks, military networks, underwater networks, and sensor networks has motivated the study and development of new delay and disruption tolerant networking protocols.

A Disruption Tolerant Network Research Group has been established with the charter to define the next generation Disruption/Delay Tolerant Networking (DTN) protocol<sup>1,2</sup>. The DTN Research Group is working on both the DTN protocol standard and software reference implementations of this protocol. The DTN Bundle Protocol is currently under development. The most significant enhancement of DTN protocols is that they operate using a store-and-forward philosophy. Therefore, internal nodes must have the ability to store incoming information and schedule a retransmitting of that information at a later time depending on the link availability. As such, information can progress through a DTN network in incremental manner as opposed to directly moving data from its origin node to the destination node all at once as is required by TCP/IP. Currently two software implementations of the DTN Bundle Protocol are under development, the Delay Tolerant Networking Research Group's DTN2, and the Jet Propulsion Laboratory's Interplanetary Overlay Network (ION). The DTN protocols have many open research areas still under investigation, such as routing algorithms, data storage and replication, reliable data transport, low complexity algorithms, low power implementations for mobile nodes, and low cost implementations<sup>3,4 and 5</sup>. Through this project Cal Poly has contributed to this fundamental research and is participating in the first implementation of a broad area DTN network.

In fall 2005 the California Central Coast Research Partnership (C3RP) provided support to establish a DTN research project at Cal Poly. This funding supported several research activities in the definition, architecture, implementation, and performance evaluation of DTN networks. The activities included: participation in the DTN Research Group, implementing a DTN network at Cal Poly, connecting the Cal Poly DTN network to the Jet Propulsion Laboratory (JPL), designing hardware support for a DTN node, and selecting a model DTN application. Currently we are seeking additional support to continue the design of hardware support for a DTN node and the implementation of a model application to demonstrate the capabilities of a DTN network.

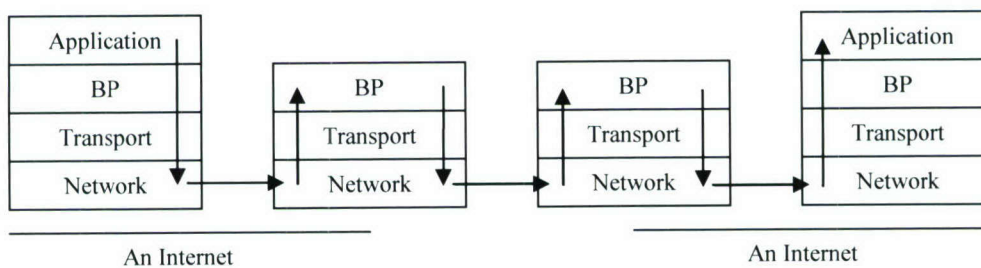
In our work, the Principal Investigators and their graduate students have established a Cal Poly DTN router that is used as the backbone for the first United States transcontinental DTN

network with JPL, the NASA Marshall Space Flight Center (MSFC), and the Johns Hopkins University Applied Physics Laboratory (APL). Our work includes the study and optimization of the software implementation and algorithms of the DTN protocol.

In addition, we are designing the first hardware support module for a DTN node using reconfigurable hardware. This hardware module will allow us to implement small mobile hand held “end” nodes that can be used in rugged non-traditional networking environments. This system on chip design would allow devices to be “DTN-enabled” without the need for an entire computer system and will conform to the DTN protocols that are currently being established. This design features pipelined hardware modules implemented in an FPGA which interface with firmware running on two embedded hardcore PowerPC processors. In tandem these systems create a DTN node, which can be used to route network traffic and/or to send and receive DTN Protocol Data Units (PDUs). This hardware-based implementation is used to further validate the bundle protocol through testing in an established ION-based DTN between Cal Poly, JPL, the NASA Marshall Space Flight Center and the Johns Hopkins Applied Physics Laboratory. Eventually our design can be integrated into an Application Specific Integrated Circuit (ASIC) or processor that is used for the mobile computing platform. More specific details of the implementation for hardware support for a DTN node can be found in the “*Delay and Disruption Tolerant Network Implementation with Reconfigurable Hardware Support*” thesis written by Don Heyer<sup>6</sup> and funded by this C3RP grant.

Currently, we are in the process of identifying a target application to demonstrate reliable data transfer with limited connectivity and unpredictable link availability in a dynamically reconfiguring network topology (i.e., underwater autonomous robotic vehicles or mobile computational nodes). The Investigators have narrowed down the potential model DTN applications to implement to one of two possibilities. We visited the Cal Poly Avila Research Pier with JPL engineers to evaluate a DTN application in May 2006. At the Avila Pier, Dr. Mark Moline and his colleagues use autonomous underwater robotic vehicles to gather ocean and marine data. Currently, these robots are not able to communicate with each other to better coordinate their tasks. Underwater communication is difficult since there are low bandwidth, limited link connectivity, and unpredictable communication link outages. The second model DTN application under consideration is a mobile computing node that could be used in a variety of applications from wireless sensor networks to battlefield communication networks.

**Overview of the DTN Bundle Protocol:** The bundle protocol is the delay tolerant equivalent to TCP. This new protocol operates at the application layer of the stack in the Open System Interconnection Basic Reference Model<sup>7</sup> (OSI Model) and moves information through the network incrementally using a store-and-forward technique.



**Figure 1: The Bundle Protocol in the OSI Model<sup>7</sup>**

The bundle protocol operates using two main types of information: data bundles and administrative record bundles. Data bundles contain the information meant to be sent and received while administrative record bundles provide control of and information about the data bundles as they move throughout a network. These bundles are important because data bundles do not (in most cases) successfully get sent from source to destination in one transmission.

As mentioned earlier, the fundamental difference between the bundle protocol and other protocols that allows it to operate in a delay tolerant network is the store-and-forward capability. Each time a bundle is successfully transmitted from one DTN node to the next, the new node assumes responsibility for transmitting the bundle to its destination (to be referred to as taking “custody” of the bundle); it essentially becomes the new source. If the bundle can immediately be forwarded to the next node in the path it is, however if any further transmission is not currently possible the bundle is stored in the node’s memory and retransmitted at the next available opportunity. As data bundles move forward throughout the network the DTN nodes generate administrative record bundles and send them back to previous nodes in the network to provide information about successful data reception and custody acceptance. Once a node receives an administrative record with information that the next node has accepted custody of a forwarded bundle, the old node may then delete the bundle from its memory.

The store-and-forward technique helps to overcome the difficulties experienced by TCP in difficult networking environments:

1. *Long and/or Variable Delays*: Long and variable delays do not cause a breakdown of the bundle protocol as they do with TCP. Whereas TCP requires that segments of data be acknowledged and that they arrive in relatively the correct order, the DTN Bundle Protocol has neither requirement for bundles of data. The DTN Bundle Protocol does implement a form of acknowledgement through the use of administrative records, however there are very relaxed bounds on the amount of time before an administrative record must propagate back through the network. If an intermediate node takes custody of a bundle, then the main disadvantage is that the intermediate node must store the data bundle in memory for an undetermined amount of time. When a data transmission link is available, the intermediate node that has taken custody of the data bundle will transmit the bundle to the next

intermediate node in the network and the protocol does not break down. Furthermore, the bundle protocol does not require that bundles arrive in order. Therefore data spread across several bundles that are transmitted over different paths, each with different delays, will not be lost.

2. *Intermittent Connectivity:* The DTN Bundle Protocol does not require a complete end-to-end connection throughout the network to transmit information from a source node to the destination node. The DTN Bundle Protocol only requires that a pathway exists between the source and destination nodes and that each hop, or link, occasionally will function. Since each node stores the bundle and waits to retransmit the bundle until the next link is available the source node may not ever need to have a complete, end-to-end path to the destination node. DTN nodes operating under the Bundle Protocol can store a bundle indefinitely until the next connection is available or until the information inside the bundle expires because it is no longer useful.
3. *High Data Error Rates:* TCP is designed to operate in environments that have low data error rates; in the rare case that an error occurs the data must be retransmitted from the original source. If a data error were to occur in a DTN Bundle Protocol network, the bundle received in error would be retransmitted from the intermediate node with custody of the bundle rather than from the original source. Therefore the bundle protocol limits the probability of a data error to the probability of an error over a single link rather than over the entire data path. While errors will typically occur more frequently in DTN environments, these errors have a smaller impact on the network because they only require the retransmission of data over a partial segment on the network rather than the entire network.

***Demonstration of the DTN Bundle Protocol:*** A DTN Bundle Protocol network has been established between JPL, the NASA Marshall Space Flight Center (MSFC), the Johns Hopkins University Applied Physics Laboratory (APL) and Cal Poly<sup>8</sup>. This network is based on JPL's "ION" DTN Bundle Protocol implementation. In this network JPL, the MSFC and the APL computers were configured as network endpoints and a DTN router was located in the NetPRL lab at Cal Poly. The DTN Bundle Protocol Network configuration as shown below:

In this network all DTN traffic created from a simple test application, *ionsource*, was inserted into a bundle by ION, then encoded in a TCP segment and sent to the internet (TCP/IP is used for bundle transmission in this network for ease of use in initial DTN development and testing). Bundles are routed to the DTN router node hosted on the netprl3 server at Cal Poly. The DTN router node then forwards the bundle to its destination node over an internet link where it is decoded by the local version of ION and passed to the test application, *ionsink*. The *ionsink* application then displays the received information on the computer screen. Proof of concept tests were run using this established DTN, starting with the first demonstration of the ION network infrastructure on March 31<sup>st</sup>, 2006<sup>8</sup>.

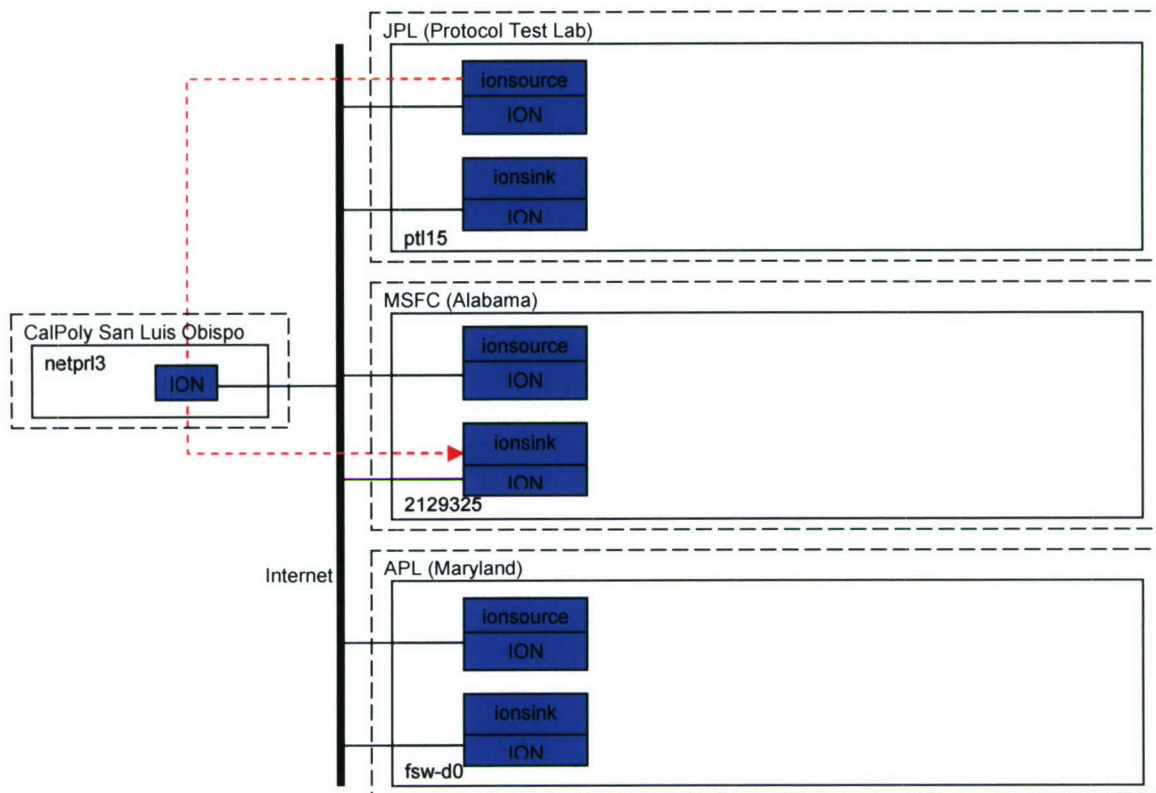


Figure 2: Established DTN Configuration<sup>8</sup>

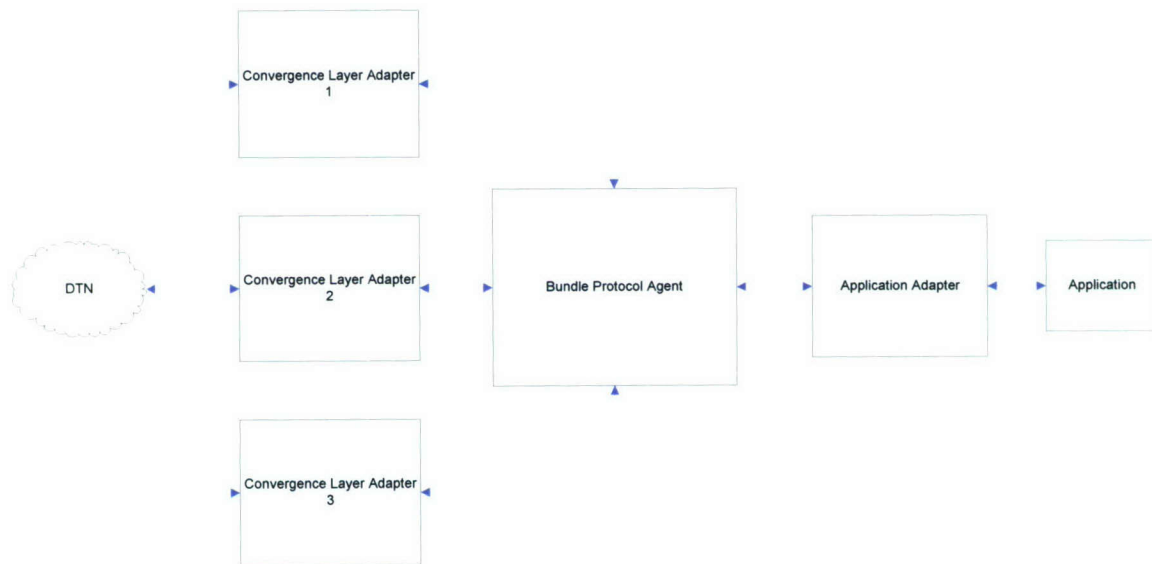
**Architecture of a DTN Bundle Protocol Node:** A DTN node is typically made up of three main components: a *convergence layer adapter*, a *bundle protocol agent*, and an *application agent*<sup>7</sup>. Figure 3 shows an example of a DTN node that links to three separate convergence layers. The specific function of each component follows:

**Bundle Protocol Agent:** The bundle protocol agent is responsible for the main processing of any bundles received or transmitted by the node. It provides the services specified by the bundle protocol by decoding, interpreting and storing incoming bundles in memory, managing any and all bundles stored in memory, and creating and forwarding bundles on to their next destination.

**Application Agent:** The application agent is the interface between any application wishing to send and receive information over the network and the bundle protocol agent. The application agent is responsible for receiving information from the application, forming new bundles for transmission, and passing any data received and decoded by the bundle protocol agent to the application. A DTN node may exist without any application agent, at which point it becomes a DTN Router with only the ability to receive and forward bundles.

**Convergence Layer:** As seen in Figure 1, the bundle protocol can operate over several different internets all running different protocols lower in the network stack. The convergence layer is the interface between the bundle protocol agent and whichever protocol

is being used for data transmission further down the network stack. For example, a bundle sent from Mars to a personal computer in somebody's home would travel over some sort of radio link to an earth satellite, where it would then be received and routed over the internet to the personal computer. Therefore the node on Mars would require a radio convergence layer to send the bundle, the node on the earth satellite receiver would require a radio convergence layer to receive the bundle and a TCP/IP convergence layer to forward the bundle over the internet, and the node on the home PC would require a TCP/IP convergence layer to receive the bundle.



**Figure 3: Overall DTN Node Structure**

**Hardware Support Implementation of the DTN Bundle Protocol:** This section includes a system level description of the reconfigurable hardware-based DTN node implementation. Figure 4 below shows the system level diagram of the DTN Bundle Protocol node.

Following the data path through this node, information is first received from the DTN Bundle Protocol Network (shown by the DTN Cloud in Figure 4) over a CAT5 Ethernet cable. The information is processed by the Ethernet MAC module running in the FPGA which forwards the information to the TCP/IP convergence layer that is implemented on the PowerPC Processor 0. The convergence layer accepts the TCP session and begins to receive TCP segments and strips the TCP overhead to produce DTN bundles. The DTN bundles are then passed up the network stack to the first stage of the bundle protocol agent pipeline, bundle decoding.

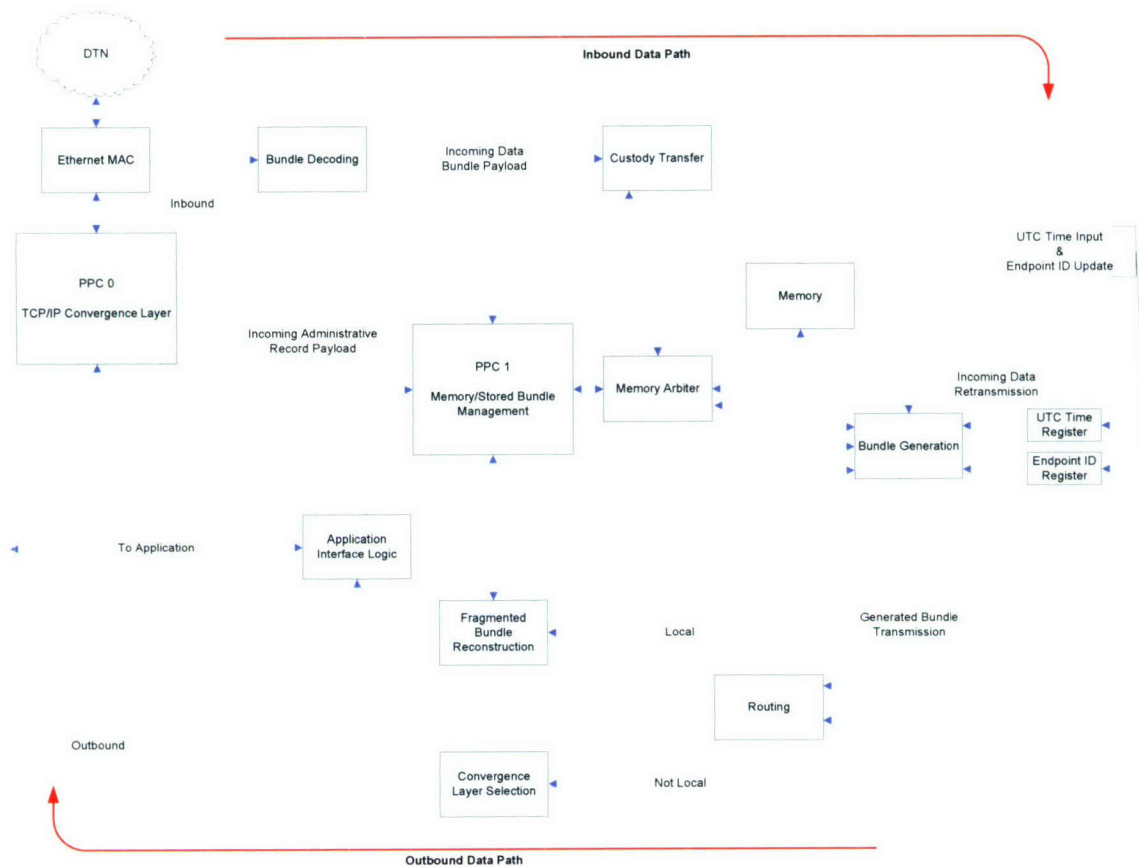
The bundle protocol agent divides the processing tasks of the bundle protocol into several pipelined stages which all interface to a central PowerPC processor which is primarily responsible for managing the node's memory storage. The bundle decoding stage reads the bundle header(s) and determines several important pieces of information about the bundle including its length and whether it is an administrative record bundle or a data bundle. In the case of an administrative record bundle the memory management processor is notified so that

it can adjust its storage accordingly, if it is a data bundle it is passed to the next stage of the pipeline, custody transfer.

The custody transfer stage of the pipeline is responsible for storing the received bundle into memory (assuming custody transfer has been requested by the source node). It requests an address from the memory management processor and writes the received bundle into memory at the proper location. Once stored into memory, the bundle is passed on to the outbound data path for forwarding to the next node in the network.

The outbound path of the pipeline consists of routing and a convergence layer selection stages. The routing stage determines the next destination node for the bundle based on its final destination, constructs a new bundle header, and passes the updated bundle to the outbound convergence layer selection stage. The convergence layer selection stage acts as a multiplexer between all of the convergence layers in the node; based on the destination node selected by the routing stage the convergence layer section stage passes the bundle to the appropriate convergence layer. There is only one convergence layer implemented in this design, this stage currently is simply a pass-thorough stage and is shown for completeness when future convergence layers are implemented. Once selected, the outbound convergence layer forwards the updated bundle to the next node in the DTN.

The application interface logic is also shown at this time for completeness; it is not yet implemented and as such this DTN node implementation behaves as a DTN router as described earlier. Each part of the node is described in more detail later in the *“Delay and Disruption Tolerant Network Implementation with Reconfigurable Hardware Support”* thesis written by Don Heyer<sup>6</sup> and funded by this C3RP grant.



**Figure 4: Node System Diagram**

**Detailed Description of work completed:** This project has established a DTN research team at Cal Poly. The principal investigator and research assistants have completed the following activities:

1. Become a participant in the existing DTN Research Group
2. Implemented a DTN router and network at Cal Poly
3. Connected the Cal Poly DTN network to JPL and other research sites
4. Studied the algorithms, performance and hardware requirements of a DTN node
5. Design and Implementation of hardware support for a DTN node using reconfigurable logic
6. Identified potential DTN applications for a technology demonstration

Currently we are seeking additional support to select a model application to demonstrate the capabilities of a DTN network with our partners.

**1. DTN Research Group:** The DTN Research Group consists of engineers and scientists from government agencies, industry, and universities. The DTN Research Group has published numerous articles related to delay and disruption tolerant networking. They are currently defining the DTN protocol and working on a software implementation of the DTN protocol. The software implementation will run on a standard personal computer and allow

nodes to be networked using the DTN protocol. The PI and research assistants have become a participant in the DTN Research group and have visited JPL quarterly during the duration of this project to participate and contribute to the DTN Research Group.

2. *Implemented a Delay Tolerant Network:* JPL has completed a first version of a software implementation of DTN Bundle Protocol called ION that runs on a personal computer. Using their implementation, the PI and research assistants have implemented a DTN node on-site at Cal Poly.

3. *Connected the Cal Poly Delay Tolerant Network to JPL:* On March 31<sup>st</sup>, 2006 the first transcontinental delay tolerant network using JPL's DTN software was established. This network was established between JPL, the NASA Marshall Space Flight Center (MSFC), Johns Hopkins University Applied Physics Laboratory (APL), and Cal Poly. This test was repeated on April 7<sup>th</sup>, 2006 and April 25<sup>th</sup>, 2006 to correct bugs that were found in the previous tests and a performance benchmark test was conducted on May 5<sup>th</sup>, 2006. In these tests, the Cal Poly software DTN node acted as a DTN router that formed the backbone for the network demonstration.

4. *Algorithm, performance, and hardware requirement study:* We have studied the performance characteristics of the network, and determined algorithmic optimizations and assessed the requirements for a hardware implementation. The results of this work will be submitted for publication and conference presentations in conjunction with later work as it is completed.

5. *Design Hardware Support for a DTN Node:* We are completing the design of a high-performance low-cost DTN node using hardware support suitable for small mobile wireless devices. This system on chip design would allow devices to be "DTN-enabled" without the need for an entire computer system and will conform to the DTN protocols that are currently being established. We used a Xilinx Virtex II Pro FPGA to design and test the DTN hardware support. Eventually our design can be integrated into an Application Specific Integrated Circuit (ASIC) or processor that is used for the mobile computing platform.

6. *Select and Model DTN Application:* The Investigators have narrowed down the potential model DTN applications to one of two possibilities. We visited the Cal Poly Avila Research Pier with JPL engineers to evaluate a DTN application in May 2006. At the Avila Pier Dr. Mark Moline and his colleagues use autonomous underwater robotic vehicles to gather ocean and marine data. Currently, these robots are not able to communicate with each other to better coordinate their tasks. Underwater communication is difficult since there are low bandwidth, limited link connectivity, and unpredictable communication link outages. The second model DTN application under consideration is a mobile computing node that could be used in a variety of applications from wireless sensor networks to battle field communication networks.

**Phase 1 Activities Complete and Schedule:** The work activities and schedule are listed in the table below.

Task	Brief Description	Original Duration	Actual Duration
1. DTN Research Group	Participate in DTN Research Group	Aug '05 – July '06	Aug '05 – Aug '06
2. Implementing a DTN Network	Using the software reference design running on PC computers.	Sept '05 – Dec '05	Sept '05 – Feb '06
3. Connect Cal Poly DTN network to JPL	System integration with other DTN networks.	Dec '05 – Jan '06	Feb '06 – Aug '06
4. Algorithm and hardware requirement study	Study algorithms and ref. implementation, propose optimizations and define high-performance, low-power improvements.	Feb '06 – March '06	Jan '06 – April '06
5. Select DTN model Application	Identify and design a model of an application that runs over a DTN network	April '06 – June '06	April '06 – Sept '06
6. DTN Hardware Support	Design hardware support for DTN node	June '06 – July '06	May '06 – Dec '06
7. Phase 1 Report		Sept '06	December '06

**Phase 2 Additional Activities and Proposed Future Funding Opportunity:**

Task	Brief Description	Approximate Duration
1. DTN Research Group Continued Collaboration	Continued Participation in DTN Research Group	Sept '06 – Sept '07
2. DTN Network Technology Demonstration	Continue to perform DTN routing for Global DTN applications	Sept '06 – Sept '07
3. Solicit NASA and other agencies for funding to support collaboration and DTN Testing	Continue to serve as the DTN backbone for network testing and technology Demonstrations	Sept '06 – Dec '06
4. Selected DTN model application and establish requirements	Implement either 1. Underwater DTN NW 2. Mobile DTN NW	Sept '06 – Dec '06
5. Implement Hardware DTN node	Design and implement HW DTN node using FPGA technology.	Sept '06 – March '07
6. Implement Complete DTN Model application	Design hardware support for DTN node	Sept '06 – Sept '07
7. Phase 2 Report		Sept '07

**Facilities:** The project has utilized the facilities of the Reconfigurable Hardware Research Lab and NetPRL lab in Building 20. This includes 4 computers that have been purchased for previous projects, 4 Xilinx XtremeDSP Development Kits (\$10k), several FGPA development boards, and digital lab equipment.

**Faculty and Student Participation:** The principle investigator and one graduate student were supported in Phase 1 of this proposal.

**Related Conferences and Meetings Attended:** The C3RP funding has allowed the research team to attend the following related conferences and meetings:

1. Quarterly meetings at JPL regarding JPL's DTN software and future plans to conduct the multi-center proof-of-concept tests (mentioned earlier).
2. "IEEE International Symposium on Low-Power and High-Speed Chips", COOL Chips IX, Yokohama, Japan, April 2006.
3. "IEEE Space Mission Challenges for Information Technology Conference", Pasadena Ca, July 2006.

***Publications:*** The C3RP funding has allowed the research team to publish and present the following work at conferences and meetings:

1. "Evaluation of the Telecommunications Protocol Processing Subsystem Using Reconfigurable Interoperable Gate Arrays", J. Pang, A. Liddicoat, P. Pingree, L. Torgerson, Proc. IEEE International Symposium on Low-Power and High-Speed Chips, COOL Chips IX, Yokohama, Japan, April 2006.

**Bibliography:**

1. Vint Cerf et. al., "Delay Tolerant Network Architecture," draft irtf from the Delay Tolerant Networking Research Group, July 2005.
2. M. Demmer, E. Brewer, K. Fall, S. Jain, M. Ho, R. Patra "Implementing Delay Tolerant Networking," Intel Research Berkeley, IRB-TR-04-020, December 2004.
3. S. Jain, K. Fall, R. Patra, "Routing in a Delay Tolerant Network," SIGCOMM, August 2004.
4. K. Fall, "Messaging in Difficult Environments," Intel Research Berkeley, IRB-TR-04-019, Dec 2004.
5. Q. Li and D. Rus, "Communication in Disconnected Ad Hoc Networks using Message Relay," Journal of Parallel and Distributed Computing, 2003.
6. D. Heyer, "Delay and Disruption Tolerant Network Implementation with Reconfigurable Hardware Support" to be published January 2007.
7. K. Scott, and S. Burleigh, "Bundle Protocol Specification", Version 7, 2006.
8. S. Burleigh, "Wide-Area Network AMS Test Report", 2006.

**Mechanistic study of the cyclin-dependent kinases PfPK5/Pfcyc-1 and PfPK6  
from the malarial parasite, *Plasmodia falciparum***

Project Investigator:

Lisa M. Lindert  
Chemistry & Biochemistry  
California Polytechnic State University  
San Luis Obispo, California

## ABSTRACT

---

Malaria is a worldwide health concern caused by the eukaryotic parasite *Plasmodia falciparum*. Sub-Saharan African children are most affected, but there is also an impact on tourists and military personnel in malaria-infested regions(1). There is a need for novel anti-malarial treatments due to increased parasite resistance to current therapies. Cyclin-dependent kinases (CDKs) have been identified as a drug target due to their essential role in regulating cell multiplication (2). A challenge to their use as anti-malarial targets is the presence of CDK homologues in the host and parasite. However, recent work has discovered important differences between human and plasmodial CDKs that may allow design of therapeutics targeted exclusively to plasmodial CDKs (3). Our study focuses on two plasmodial CDKs, PfPK5/Pfcyc-1 and PfPK6. Our goal is to determine the catalytic mechanism of these malarial CDKs for comparison with each other and the well-studied human CDK2/cyclin A. PfPK5/Pfcyc-1 is similar to the well-studied human CDK2/cyclin A and PfPK6 is closely related to PfPK5, but does not require activation by an associated cyclin subunit (Pfcyc-1). Recombinant PfPK5/Pfcyc-1 and PfPK6 have been produced and purified in quantities sufficient for enzymatic study in our laboratory at Cal Poly. Both preparations have been assayed for enzymatic activity with respect to a common CDK substrate histone H1 and have displayed differing kinetic properties from each other and from CDK2/cyclinA. Our current goal is to characterize the kinetics of substrate binding and catalytic turnover with respect to synthetic peptide substrates and adenosine triphosphate (ATP). This will be determined *in vitro* using a spectrophotometric assay. This research will be performed in collaboration with a group of malarial researchers worldwide dedicated to the design of anti-malarial therapeutics. Our research at Cal Poly will be a key part of this investigation and provide undergraduates the opportunity to learn basic biochemical techniques and see their work directly applied to a global problem.

## SIGNIFICANCE

---

Malaria affects over 400 million people each year, mostly in sub-Saharan Africa. It has its largest impact on small children, but also impacts tourists and military personnel in that region. The largest challenge to treatment is plasmodial resistance to current therapeutics (1). To combat this, new targets for anti-malarial therapies are being explored. Basic research into the life cycle of the malarial parasite and publication of the *P. falciparum* genome have been major steps(4). These efforts have resulted in the identification of genes encoding plasmodial CDKs and several plasmodial cyclins(1,3). This research is moving toward the development of CDK-targeted anti-malarial compounds based on the structures of available anti-cancer agents developed to target mammalian CDKs. The goal is the development of therapeutics specific to malarial CDKs. This is a feasible goal based on the known structural and functional differences observed between plasmodial and human CDKs(1). Currently, our collaborators are investigating the molecular structure of plasmodial CDKs to identify key structural differences between plasmodial and human CDKs and screening novel compounds to identify new inhibitors(1,3,5,6). These researchers reside in academia and the US military. They include Dr. Jane Endicott, from the University of Oxford, Dr. Christian Doerig, from the Wellcome Center for Molecular Parasitology at the University of Glasgow and Dr. Norman Waters, from the Walter Reed Army Institute of Research in Maryland. The focus of this proposed research is a functional analysis of plasmodial CDKs that will lead to rational drug development.

## SUMMARY OF PROGRESS TO DATE

---

Our progress toward the development of an undergraduate research program in the study of malarial CDKs has progressed significantly. Since the funding of the original C3RP and CSUPERB grants on PfPK5 and PfPK6, a total of sixteen undergraduate students have worked on this project at various stages and levels. Several students working this project have presented their work as posters at the American Chemical Society meetings in Anaheim (March 2004) and Philadelphia (August 2004), a CSUPERB meeting (January 2005), and the Centennial National Meeting of the American Society of Biochemistry and Molecular Biology in San Francisco (April 2006). Eight of these students have graduated and they have gone on to work in California biotech companies, enrolled in medical school or have gone to graduate school. Four of the continuing students on the project will apply for Ph.D. programs for admission in 2007.

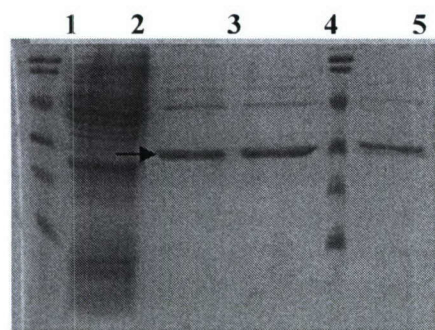
There has been significant progress toward the research goals of this project as well. During the summer months of 2004 students worked to refine the protocol we had developed for purification of the PfPK5 and cyc-1 proteins as recombinant GST-fusions in *E. coli* (Figure 1). The protein was produced in sufficient quantities to allow enzymatic analysis with respect to a small protein histone H1. In addition, students working on PfPK6 have also successfully expressed this protein in *E. coli* and purified it from cell lysates. The results are shown in Figure 2, below.

My lab has produced large quantities of PfPK6 and we are working to produce large amounts of PfPK5/cyc1 for enzymatic analysis. We have assayed preparations of each enzyme and found that both demonstrate activity lower than that expected for these enzymes. The activity of the enzyme we purified has varied between preparations suggesting a difference in phosphorylation state or CDK-cyclin binding.

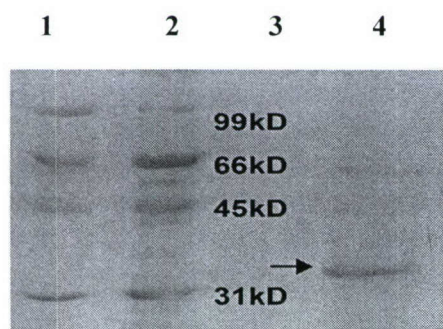
Because PfPK5 autophosphorylates after cyclin binding, this is a particular problem with this preparation. We believe this is due to presence of unequal amounts of folded PfPK5 and cyc-1 in our preparations. This is difficult to assess due to the similarities in the molecular masses or the recombinant protein constructs we currently use. We are working to produce constructs with His<sub>6</sub> tags which will allow us to perform reciprocal pull-down experiments to test this hypothesis and validate our preparations. We are currently assessing the phosphorylation state of the enzyme to determine whether it has been activated by autophosphorylation using a phosphoprotein stain (Pro-Q Diamond, Invitrogen). Previous reports suggested that this process occurs during expression and purification in *E. coli* however we have not assessed this in our lab (14,16) and see this as a key control that we can now perform due to the availability of this reagent and a fluorescence gel imager on our campus.

Preliminary data obtained by Kristina Moody in 2005-2006 suggests that PfPK6 has substrate specificity that is different from CDK2 and more similar to another family of kinases the MAP kinases. This is a very significant finding and is substantiated by work from our collaborators at Oxford (J. Endicott, personal communication). We will be actively pursuing this avenue through the use of synthetic peptide substrates of varying sequence and through work with our collaborators. Both substrates are phosphorylated differently, but poorly in our assays. This indicates a challenge for our work in finding a peptide substrate, but an opportunity as well. This shows a significant difference from the human CDK and MAPK enzymes. The CDKs selectively phosphorylate histone-like substrates very well and MAPKs phosphorylate MBP-like

substrates very well. It is very unique to have a protein that is more CDK-like phosphorylate an MBP substrate. Analysis of the autophosphorylation sites in combination with these kinetic data will allow us to better predict an optimal PfPK6 substrate. This will aid inhibitor design and the identification of *in vivo* substrates in the *P. falciparum* proteome.



**Figure 1: SDS-PAGE analysis of PfPK5/cyc-1 purification.** Lanes 1 and 5 contain protein molecular weight standards. The sizes of the bands from top to bottom are 103,77,50,34.3,28.8, 20.7 kDa. Lane 1 is the bacterial cell lysate sample containing GST-PfPK5 and GST-cyc1. Lanes 3, 4 and 6 contain fractions eluted following protease cleavage (Prescission Protease, Amersham Biosciences) of the PfPK5/PfPKcyc-1 complex from its GST-tag on the glutathione agarose column. PfPK5 and cyc-1 are very similar in size, ~34kDa. The band corresponding to the putative PfPK5/PfPKcyc-1 is indicated with an arrow. HPLC analysis of non-denatured PfPK5/ confirmed the presence of a protein with the expected size of the PfPK5/cyc-1 heterodimer, 79.5 kDa (data not shown).



**Figure 2: SDS-PAGE analysis of PfPPfPK6 purification.** Lanes 1 and 2 contain protein molecular weight standards. The sizes of the bands from top to bottom are 99, 66, 45, 31 kDa and are shown on the gel. Lane 3 is blank. Lane 4 contains purified PfPK6 cleaved from its GST-Tag on the glutathione agarose column (Prescission Protease, Amersham Biosciences). PfPK6 has a ~36kDa and is shown with an arrow in lane 4.

## REFERENCES

1. Waters, N. C., and Geyer, J. A. (2003) *Expert Opin Ther Targets* **7**, 7-17
2. Doerig, C., Chakrabarti, D., Kappes, B., and Matthews, K. (2000) *Prog Cell Cycle Res* **4**, 163-183
3. Doerig, C., Endicott, J., and Chakrabarti, D. (2002) *Int J Parasitol* **32**, 1575-1585
4. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M. S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M., and Barrell, B. (2002) *Nature* **419**, 498-511
5. Doerig, C., Meijer, L., and Mottram, J. C. (2002) *Trends Parasitol* **18**, 366-371
6. Holton, S., Merckx, A., Burgess, D., Doerig, C., Noble, M., and Endicott, J. (2003) *Structure (Camb)* **11**, 1329-1337
7. Morgan, D. O. (1997) *Annu Rev Cell Dev Biol* **13**, 261-291

8. Stevenson, L. M., Deal, M. S., Hagopian, J. C., and Lew, J. (2002) *Biochemistry* **41**, 8528-8534
9. Hagopian, J. C., Kirtley, M. P., Stevenson, L. M., Gergis, R. M., Russo, A. A., Pavletich, N. P., Parsons, S. M., and Lew, J. (2001) *J Biol Chem* **276**, 275-280
10. Pavletich, N. P. (1999) *J Mol Bio* **287**, 821-828
11. Johnson, L. N., De Moliner, E., Brown, N. R., Song, H., Barford, D., Endicott, J. A., and Noble, M. E. (2002) *Pharmacol Ther* **93**, 113-124
12. Haesslein, J. L., and Jullian, N. (2002) *Curr Top Med Chem* **2**, 1037-1050
13. Graeser, R., Wernli, B., Franklin, R. M., and Kappes, B. (1996) *Mol Biochem Parasitol* **82**, 37-49
14. Le Roch, K., Sestier, C., Dorin, D., Waters, N., Kappes, B., Chakrabarti, D., Meijer, L., and Doerig, C. (2000) *J Biol Chem* **275**, 8952-8958
15. Prowse, C. N., and Lew, J. (2001) *J Biol Chem* **276**, 99-103
16. Stevenson-Lindert, L. M., Fowler, P., and Lew, J. (2003) *J Biol Chem* **278**, 50956-50960
17. Adams, J. A. (2001) *Chem Rev* **101**, 2271-2290
18. Bracchi-Ricard, V., Barik, S., Delvecchio, C., Doerig, C., Chakrabarti, R., and Chakrabarti, D. (2000) *Biochem J* **347 Pt 1**, 255-263
19. Lew, J. (2003) *Biochemistry*. **42(4)**, 849-856

**Development of a parallel computational framework for computational fluid  
dynamic simulations**

Project Investigator:

David D. Marshall  
Aerospace Engineering  
California Polytechnic State University  
San Luis Obispo, California

## ***Overview***

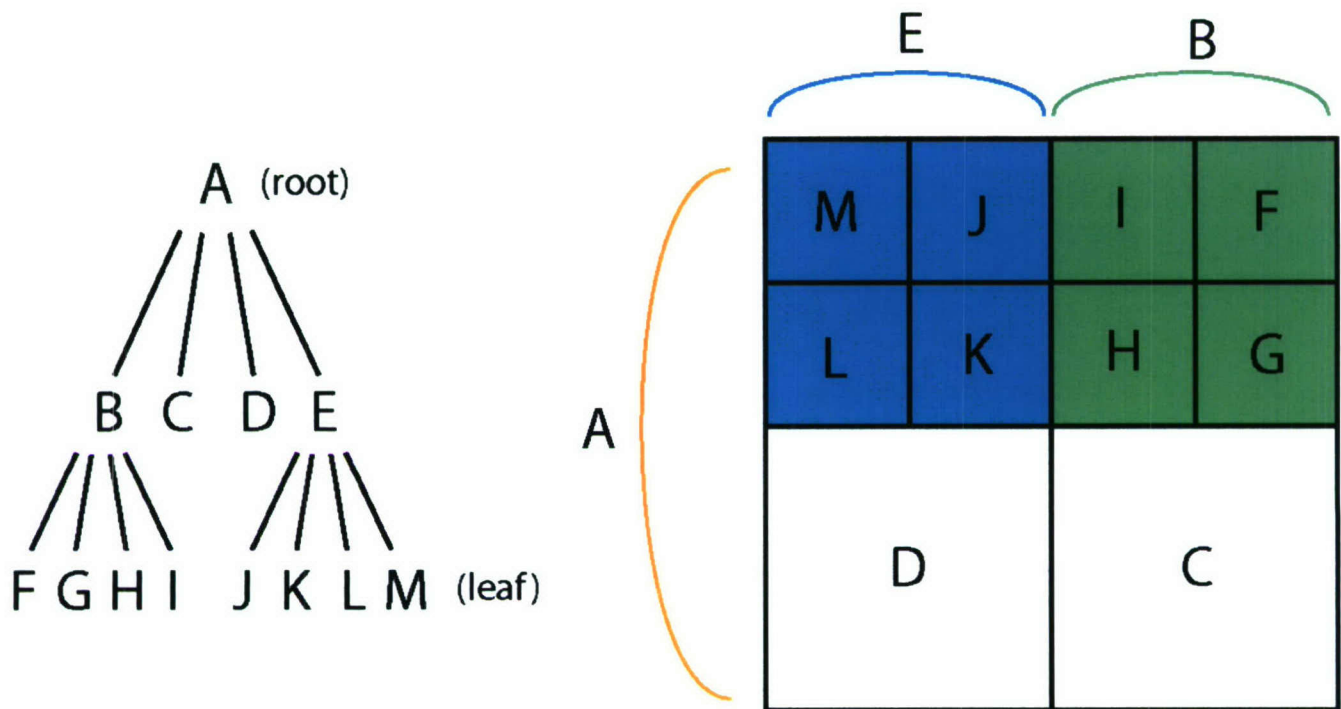
Significant work has been accomplished towards the development of the parallel computational framework. The grid generation task has resulted in a robust Cartesian grid generation code that is capable of generating high quality grids around complex geometries. The computational fluid dynamics task has developed a novel variable extrapolation technique for non-uniform grids that can result in up to third order accurate interpolations. In addition a large number of inviscid flux formulations have been implemented to be integrated into the framework.

## ***Grid Generation***

The majority of the research up to this point has been spent on creating a suitable grid generator. The objectives were to create a grid generator that would: work with both particle and CFD codes, be robust enough to handle complicated geometries, and be parallelizable. At this point, a two-dimensional, Cartesian unstructured grid generator has been developed. This generator uses a cut-cell embedded geometry approach to capturing complicated topologies. The software is object oriented and was developed using the C++ programming language.

The grid generator is based around the quadtree data structure. This data structure begins with a root data element that contains links to its four child data elements. These four data elements contain links to their parent (the root data element) and to their children. The quadtree structure was a natural choice for a Cartesian unstructured grid generator, as each cell can be divided into four equally sized “children” cells for each level of refinement. Figure 1 depicts the quadtree data structure and how it is utilized for the Cartesian unstructured grid cells.

As Figure 1 illustrates, each level of the quadtree corresponds to a refinement level in the computational grid. The cells at the bottom of the quadtree are known as leaf cells. These are the cells that will actually be used for computation by the numerical schemes. Besides being an intuitive method for organizing the grid, the quadtree data structure also provides for rapid cell searching. Since particle-in-cell (PIC) schemes require each particle to be located after every time-step, quick cell searching is essential. With a quadtree data structure, identifying neighboring cells and performing top-down searches of the tree is simple and fast.



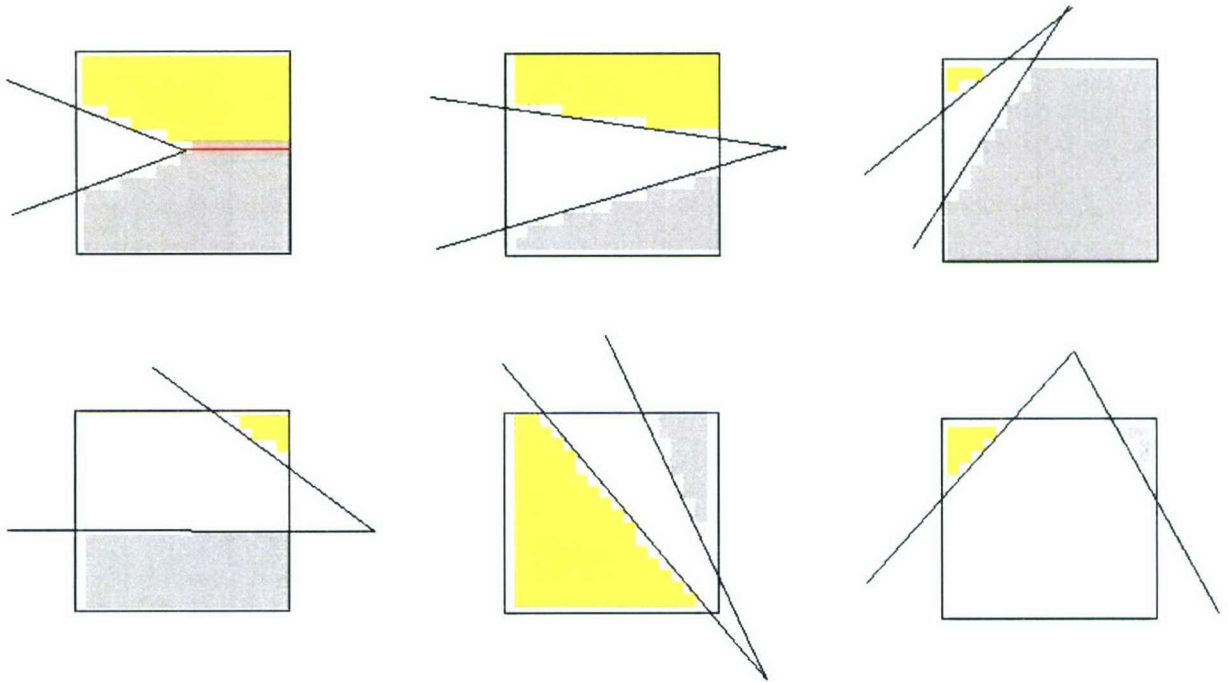
*Figure 1: Quadtree Data Structure*

The grid construction process begins with either a single root cell or a structured Cartesian grid that consists of root cells. The computational state of each root cell is then determined. That is, each root cell is set as being one of three states: cut by the geometry, in the flow field and thus computational, or non-computational. Before the root cells are refined, “boundary cells” are created around the outside of the computational domain. The boundary cells simply mirror the outermost cells. These cells are used to enforce boundary conditions for both the CFD and particle schemes. They will also be utilized as a “buffer” between processor domains once the code has been parallelized. For the particle scheme, they ensure that no particles are lost when leaving the computational and processor domains.

Once the boundary cells are created, the grid is refined according to the user specified minimum and maximum levels of refinement. Only those cells that intersect the geometry are refined to the maximum level. Every time a cell is refined, the grid is checked for “smoothness.” That is, a 2:1 ratio is enforced between adjacent cells. In other words, cells that share a side must be at most one level away from each other in the quadtree. This smoothness is important, because it ensures that there won’t be any major changes in cell size, which could cause computational instability and errors. When a cell is refined, its neighbors are checked for the 2:1 ratio, and if required, they are refined. If they are refined, then they are also checked for smoothness. This recursive pattern ensures that the entire grid is smooth throughout the entire process.

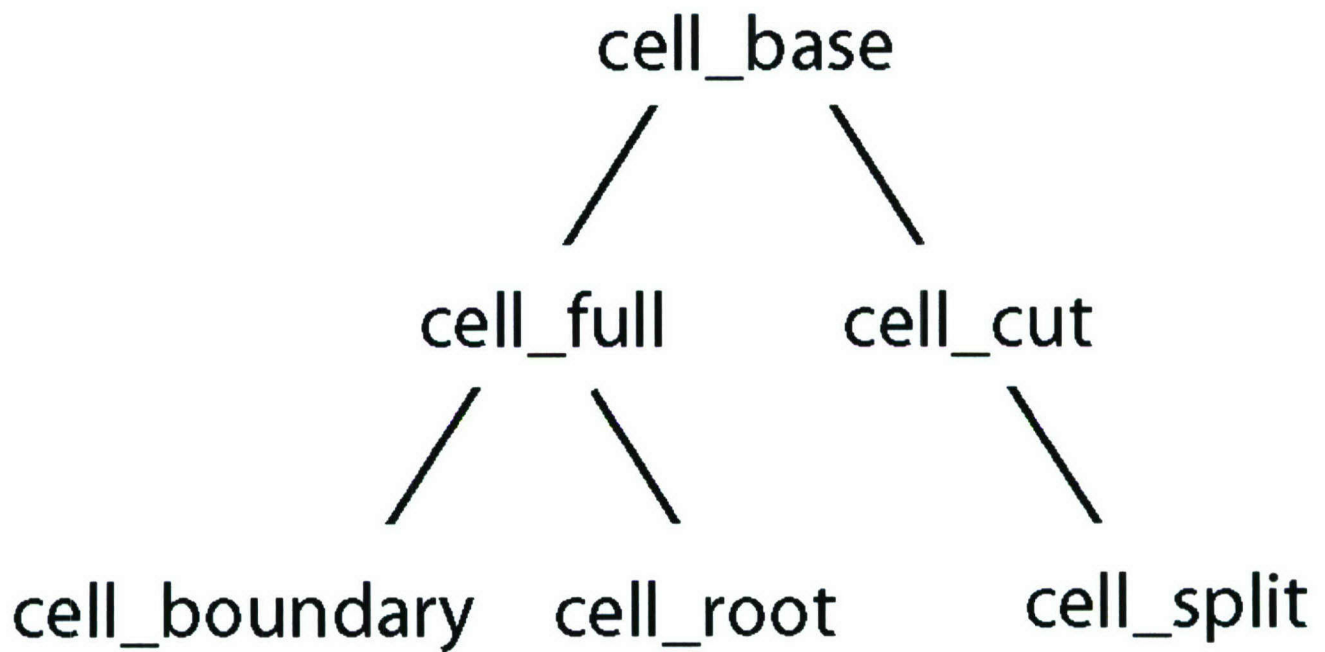
When a cell that intersects the geometry has reached the maximum level of refinement, it is turned into a cut cell. The cut cells are computed using a polygon clipping algorithm based on the one polygon clipping technique by Sutherland and Hodgman<sup>1</sup>, which takes advantage of the convex nature of the clipping cells. As each cell is cut, a simple crossproduct is used to determine which side of the cell to keep. The cell cutting process becomes extremely difficult

and computationally intensive as the geometry's complexity rises and the number of segments cutting each cell increases. This is further complicated by degenerative cases where the geometry cuts the cell in such a way that it causes the basic cell clipping algorithm to break down. Figure 2 depicts just a few of these cases involving two cutting segments.



**Figure 2: Cut Cell Problem Case**

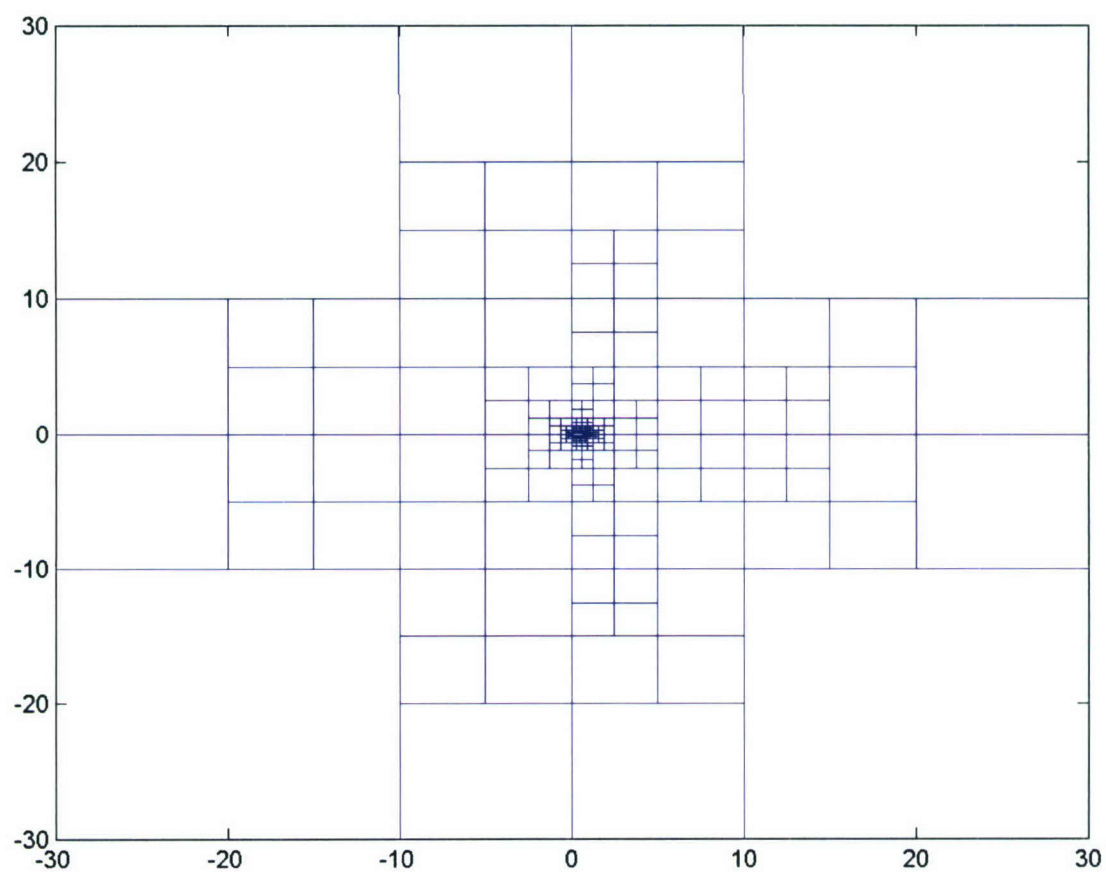
Figure 2 illustrates, many of the situations required that the cells be split into two cells, rather than just one singular cut cell. These cells are called “split cells” and simply act as two children leaf cells to a parent cell. Figure 3 depicts the basic inheritance of the different cell types that exist in the quadtree.



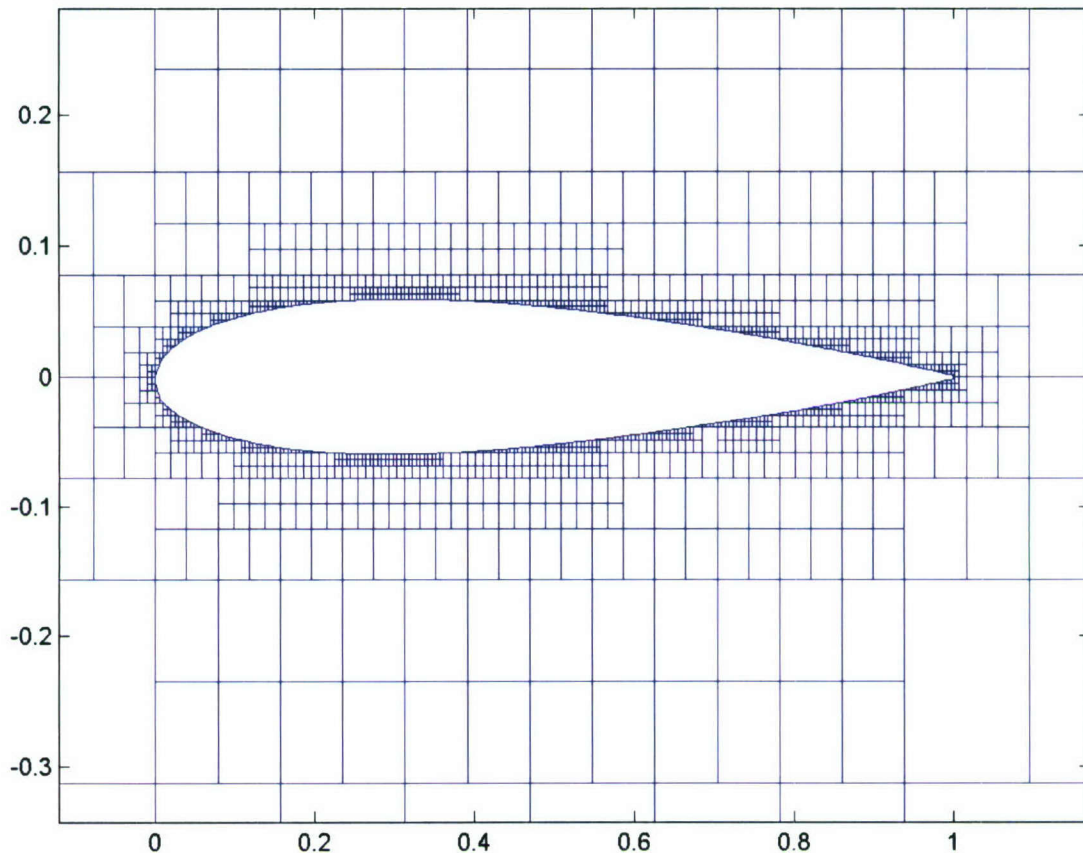
*Figure 3: Cell Class Inheritance Tree*

As Figure 3 shows, split cells are just considered a specialized cut cell. The algorithm developed will only handle cases involving two segments cutting the cell. Further work needs to be done to increase its robustness so that it can handle a wider range of geometries.

Once the cut cells are computed the grid is in a finished state. Throughout the refinement process, the aforementioned computational state of each cell is determined. The root cells that are determined to be computational or non-computational simply pass their status down to their children as they are refined. When cells that have a 'cut' computational state are refined, their children's computational states have to be determined. These computational states allow the particle and CFD schemes to access the appropriate cells. Figures 4 and 5 are full and zoomed in views of a grid generated around a NACA 0012 airfoil.



*Figure 4: Unstructured Cartesian Grid Around NACA 0012 Airfoil*



***Figure 5: Unstructured Cartesian Grid Near Surface of NACA 0012 Airfoil***

The airfoil shown in Figures 4 and 5 is composed of 99 segments. The grid has twelve levels of refinement. These figures clearly illustrate the cell cutting and smoothing algorithms previously mentioned. In addition to the already proposed future work, the grid generator should be able to handle intersections with surfaces other than simple line segments. Long-term future work might include integrating it with a CAD package. This would place all the heavy computational geometry on the CAD program and allow for the rendering of more complex geometries. An extension to three dimensions is also a longterm goal. The quadtree structure will also lend itself well to the addition of parallel processing as well as solution adaption for the particle and CFD schemes.

### ***Non-Uniform MUSCL Extrapolation***

This document is intended to provide a summary of an extension of the standard MUSCL extrapolation scheme<sup>2</sup> to be used on non-uniform grids. The use of Non-Uniform MUSCL Extrapolation (NUME) schemes on two types of grid storage schemes is possible. The two types of non-uniform grids are node-based schemes (NUME-N) where the data is stored on each node of the grid and cell-based schemes where the data is stored on the centroid of each cell (NUME-

C). These two difference storage schemes yield different NUME schemes because of the location of the resulting flux calculations. The node-based schemes assume the flux is to be calculated at the mid-point of the nodes. The cell-based schemes evaluate the fluxes at the faces of the cells which will not be at the mid-point of the cell centroids. Instead, a cell's faces are an equal distance from its centroid. In addition to the standard NUME schemes, limiters are added to the schemes using the non-uniform limiters from Berger, Aftosmis, and Murman.<sup>3</sup>. These schemes retain the overall order of accuracy for smooth function even for severe grid spacing, and appropriately limit discontinuous data. The result of this effort are two schemes that can extrapolate data on non-uniform grids with up to 3<sup>rd</sup> order accuracy that uses limiters to ensure the monotonicity of the data.

### Review of MUSCL Extrapolation

While this is a standard scheme found in any number of references, it does form the foundation for this work, so a brief review will be presented. The grid is shown in Figure 6. Since the grid is uniform, the spacing between each node is  $\Delta x$ .

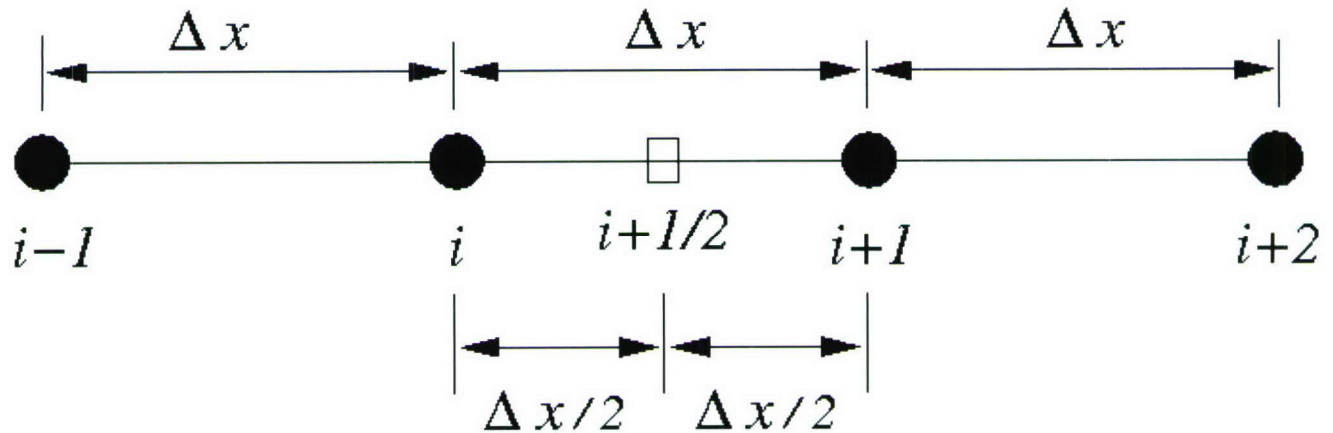


Figure 6: Uniform Grid Geometry

extrapolation functions

$$U_{i+1/2}^L = u_i \left[ \frac{1}{4} \left( 1 - \frac{\Delta x}{\Delta x} u_{i-1} \right) \right]$$

$$U_{i+1/2}^R = u_{i+1} \left[ \frac{1}{4} \left( 1 - \frac{\Delta x}{\Delta x} u_{i+1} \right) \right]$$

where the following table shows what parameters result in various schemes

Scheme Name	Order	Coefficients
Low Order	1 <sup>st</sup>	$\alpha \neq 1$ for any $\beta$
Backward	2 <sup>nd</sup>	$\alpha = 1$ and $\beta = -1$
Fromm	2 <sup>nd</sup>	$\alpha = 1$ and $\beta = 0$
Central	2 <sup>nd</sup>	$\alpha = 1$ and $\beta = 1$
Higher Order	3 <sup>rd</sup>	$\alpha = 1$ and $\beta = 1/3$

Table 1: Uniform MUSCL Coefficients

### Slope Limiters Applied to MUSCL

Slope limiters are used to preserve monotonicity near sharp gradients in the data. Applying them to the MUSCL scheme yields

$$U_{i+1/2}^L = u_i \left[ \frac{\Delta x_i}{4} \left( 1 - \lim_{\Delta x_i} \frac{u_i - u_{i-1}}{\Delta x_i} \right) \right]$$

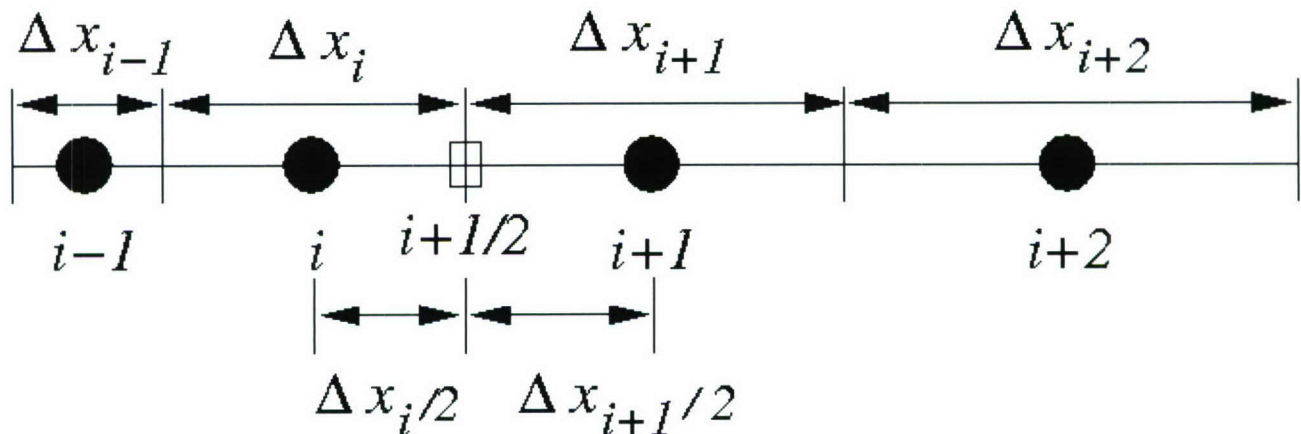
$$U_{i+1/2}^R = u_{i+1} \left[ \frac{\Delta x_{i+1}}{4} \left( 1 - \lim_{\Delta x_{i+1}} \frac{u_{i+1} - u_i}{\Delta x_{i+1}} \right) \right]$$

where  $\lim_{\Delta x}$  is the slope limiter and the limiter may be different for the left or right extrapolation. Any standard slope limiter may be used as long as it has the symmetry property that

$$\lim_{\Delta x} \frac{u_i - u_{i-1}}{\Delta x_i} = - \lim_{\Delta x} \frac{u_{i+1} - u_i}{\Delta x_{i+1}}$$

### NUME-C

This section extends the standard MUSCL scheme to be used on grids that are non-uniform with the state information being stored in a cell centered basis, cell centered NUME or NUME-C. The grid is shown in Figure 7. In this section the cell size is represented by  $\Delta x$ . Notice that the cell face  $i+1/2$ , in general, is not at the mid-point of its neighboring points.



**Figure 7: Non-Uniform Grid Geometry with Cell Centered Storage Scheme**  

$$U_{i+1/2}^{\pm} = u_{i+1/2} \left[ \frac{\Delta x_{i+1/2}}{4} \left( 1 \pm \lim_{\Delta x_{i+1/2}} \frac{u_{i+1} - u_i}{\Delta x_{i+1/2}} \right) \right]$$

where  $\lim_{\Delta x}$  is the slope limiters from Berger et al. and is a function of the state differences and the grid spacings. The following table shows what parameters result in various schemes

Scheme Name	Order	Coefficients
Low Order	1 <sup>st</sup>	$\lim_{\Delta x} \neq 1$ for any $\lim_{\Delta x}^L$ and $\lim_{\Delta x}^R$
Backward	2 <sup>nd</sup>	$\lim_{\Delta x} = 1$ and $\lim_{\Delta x}^L = -1$ and $\lim_{\Delta x}^R = -1$
Fromm	2 <sup>nd</sup>	$\lim_{\Delta x} = 1$ and $\lim_{\Delta x}^L = 0$ and $\lim_{\Delta x}^R = 0$

<i>Scheme Name</i>	<i>Order</i>	<i>Coefficients</i>
Central	2 <sup>nd</sup>	$\alpha = 1$ and $\beta^L = 1$ and $\beta^R = 1$
Higher Order	3 <sup>rd</sup>	$\alpha = 1$ and $\beta^L = \frac{\Delta x_{i-1} \Delta x_i - \Delta x_{i+1} \Delta x_{i+2}}{\Delta x_{i-1} \Delta x_i + \Delta x_{i+1} \Delta x_{i+2}}$ and $\beta^R = \frac{-\Delta x_i \Delta x_{i+1} + \Delta x_{i+1} \Delta x_{i+2}}{\Delta x_i \Delta x_{i+1} + \Delta x_{i+1} \Delta x_{i+2}}$

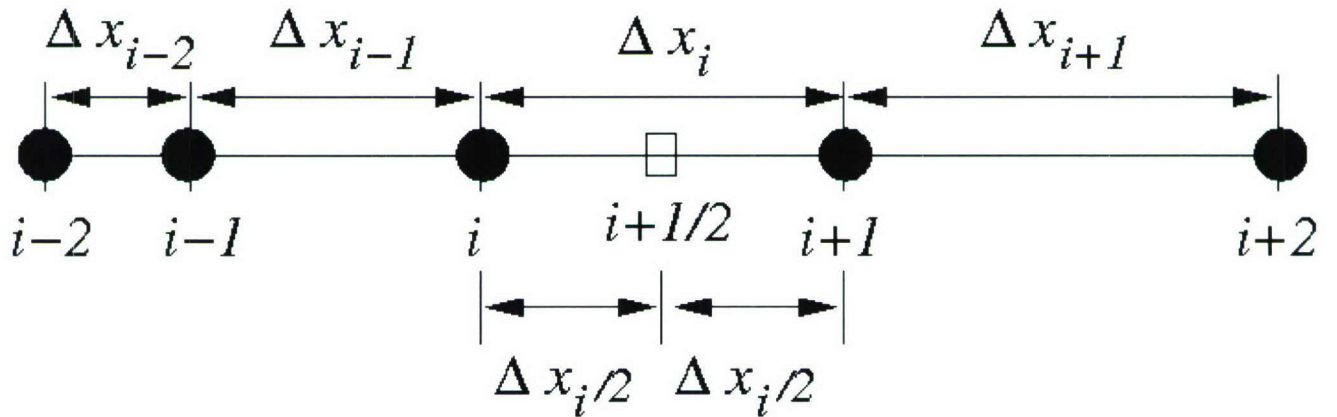
**Table 2: Cell Centered Non-Uniform MUSCL Extrapolation Coefficients**

Notice that the only coefficients that changed from the uniform formulation was the higher order scheme coefficients. Also, for a uniform grid this degenerates to the uniform formulation.

A number of numerical tests on the order and accuracy of this scheme have shown that for even extreme stretched grids where stencil used for the extrapolation had cells with sizes varying by a factor of 20 second and third order convergence characteristics are obtained from the above equations with the error in the extrapolation and slope limiting comparable to the error on a uniform grid.

#### NUME-N

This section extends the standard MUSCL scheme to be used on grids that are non-uniform with the state information being stored on the grid nodes, node based NUME or NUME-N. The grid is shown in Figure 8. In this section  $\Delta x$  is the distance from node  $i$  to node  $i+1$ . Notice that in this configuration the cell face  $i+1/2$ , is at the mid-point of its neighboring points, unlike the NUME-C scheme.



**Figure 8: Non-Uniform Grid Geometry with Node Based Storage Scheme**

$$U_{i+1/2}^R = u_{i+1} \frac{\beta^R}{4} \left\{ 1 - \beta^R \left[ \frac{2 \Delta x_i \Delta x_{i+1}}{\Delta x_i \Delta x_{i+1} + \Delta x_{i+1} \Delta x_{i+2}} \right] u_{i+1} + \beta^R \left[ \frac{2 \Delta x_i \Delta x_{i+1}}{\Delta x_{i-1} \Delta x_i + \Delta x_{i+1} \Delta x_{i+2}} \right] u_i \right\}$$

where  $\beta$  is the slope limiters from Berger et al. The limiters from Berger et al. are based on the cell volume on the left and right side of the half-node which is different from the  $\Delta x$  defined for this grid. The following table shows what parameters result in various schemes

<i>Scheme Name</i>	<i>Order</i>	<i>Coefficients</i>
--------------------	--------------	---------------------

<i>Scheme Name</i>	<i>Order</i>	<i>Coefficients</i>
Low Order	1 <sup>st</sup>	$\alpha \neq 1$ for any $\alpha^L$ and $\alpha^R$
Backward	2 <sup>nd</sup>	$\alpha = 1$ and $\alpha^L = -1$ and $\alpha^R = -1$
Fromm	2 <sup>nd</sup>	$\alpha = 1$ and $\alpha^L = 0$ and $\alpha^R = 0$
Central	2 <sup>nd</sup>	$\alpha = 1$ and $\alpha^L = 1$ and $\alpha^R = 1$
Higher Order	3 <sup>rd</sup>	$\alpha = 1$ and $\alpha^L = \frac{\alpha x_{i-2} \alpha^2 x_{i-1} - \alpha x_{i-1}}{\alpha x_{i-2} \alpha^2 x_{i-1} \alpha^2 x_i \alpha x_{i+1} - \alpha x_{i-1} \alpha^2 x_{i+1} \alpha x_{i+2}}$ and $\alpha^R = \frac{-\alpha x_{i-1} \alpha^2 x_{i+1} \alpha x_{i+2}}{\alpha x_{i-1} \alpha^2 x_i \alpha^2 x_{i+1} \alpha x_{i+2}}$

**Table 3: Cell Centered Non-Uniform MUSCL Extrapolation Coefficients**

Notice that the only coefficients that changed from the uniform formulation was the higher order scheme coefficients. Also, for a uniform grid this degenerates to the uniform formulation.

A number of numerical tests on the order and accuracy of this scheme have shown that for even extreme stretched grids where stencil used for the extrapolation had cells with sizes varying by a factor of 20 second and third order convergence characteristics are obtained from the above equations with the error in the extrapolation and slope limiting comparable to the error on a uniform grid.

### **Flux Formulations**

Another needed core component of the architecture that has been developed is the inviscid flux calculations for the CFD solver. This builds on the NUME work by using the extrapolated data to calculate the inviscid fluxes at the cell surfaces. A number of flux formulations have been developed that include:

- van Leer Flux Vector Splitting<sup>4</sup>
- AUSM Flux Vector Splitting<sup>5</sup> with both linear and quadratic pressure reconstruction
- Kinetic Flux Vector Splitting<sup>6, 7</sup> using a Maxwellian velocity distribution function
- Godunov Riemann Solver<sup>8</sup>
- Roe's Approximate Riemann Solver<sup>9</sup>
- HLLC Approximate Riemann Solver<sup>10</sup>

Each flux implementation has been tested against 5 standard and very challenging shock tube problems that have known solutions in the form of the general Riemann relations.

### **Miscellaneous Accomplishments**

All of the desired computers (three in total with additional external funds provided to purchase the third) have been purchased with the following specifications:

- Dual 64-bit Intel XEON 3.00 GHz Processors

- 4GB RAM
- 250 GB Hard disk
- 10/100/1000 MB Ethernet

Notice that these specifications exceed the proposed specifications. These machines are in their final stages of configuration. The final remaining configuration step is the testing of the parallelization functionality on these machines.

To gain an in depth understanding of the current state of the art in particle and continuum modeling, the graduate student and PI attended the AIAA Plasmadynamics and Lasers Conference that was jointly held with the Applied Aerodynamics, Fluid Dynamics, and other conferences. This has proved to be invaluable opportunity to meet with several experts in the fields of parallel computing, CFD, and particle modeling.

### ***Future Work***

#### *Topology*

In order to make this effort more useful in the future, it needs to be extended to more complicated topologies. This includes other types of computational grids beyond the two-dimensional Cartesian cells currently being developed. Also, more complex surface types need to be included such as curves and arcs. One more interesting extension is coupling the grid generation with the CAD system that is providing the surface model in order to refine the grid on the surface while keeping the original topology. This is an active area of research in the community that has not been adequately resolved.

#### *Particle Code*

A small portion of research has been done in developing a particle-in-cell (PIC) scheme to use with this grid generation. A leap-frog method has been developed for the propagation of particles. Particle searching algorithms have also been implemented. Currently, algorithms to model the interaction between particles and the surfaces of the geometry are being developed. Once these interactions are handled, DSMC methods will be implemented to model particle collisions. The long term goal is to create a parallelized PIC code that can handle collisions as well as charged particles.

Extending the continuum and particle models to handle more complex physics is another needed future effort. For the continuum model this includes adding viscous flux modeling, extending the fluid models to equilibrium air, adding chemical reactions, and coupling the particle model to the continuum model (such as mass, momentum, and energy exchange). For the particle model this includes adding collision modeling, handle charged particles, and coupling the continuum model to the particle model.

#### *Computational Fluid Dynamics Modeling*

One particular area of future is to collaborate with Berger, Aftosmis, and Murman to extend the

NUME scheme in conjunction with their irregular grid work<sup>3</sup>. Initial contact has been made with these researchers, and there appears to be significant interest in future collaboration.

Another interesting area of future work is in developing techniques to model flows that have mixed rarefied and continuum regimes. One application of this is atmospheric entry vehicles where the freestream density is too low for the flow to be considered a continuum, however after the bow shock there is a significant region of continuum flow because of the density increase from the shock. This architecture is well suited to handle this since both types of modeling are already integrated together. In addition, the Kinetic Flux Vector Splitting scheme is also well suited for the continuum region since it is based on the Boltzmann equations with a Maxwellian velocity distribution. Another continuum flux model that might be worth exploring is the Gas-Kinetic BGK scheme<sup>11</sup>. This is based on the Chapman-Enskog expansion of the BGK model for gases that are not in equilibrium (i.e. do not have a Maxwellian velocity distribution function), but are close to the Maxwellian distribution.

Finally, a number of major governmental research agencies (Department of Energy, NASA, Air Force, Navy, and Army) are all interested in the development of parallel CFD work with mixed Eulerian-Lagrangian frameworks like this.

- 
- 1 J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics, Principles and Practice*. Addison-Wesley Systems Programming Series, Addison-Wesley, 2<sup>nd</sup> edition, 1991.
  - 2 B. van Leer. Toward the ultimate conservative difference scheme. V. A second order sequel to Godunov's method. *Journal of Computational Physics*, 32(1):101-136, 1979.
  - 3 M. Berger, M. J. Aftosmis, S. M. Murman. Analysis of slope limiters on irregular grids. In *43<sup>rd</sup> AIAA Aerospace Sciences Meeting and Exhibit*, AIAA-2005-0490, 2005.
  - 4 B. van Leer. Flux vector splitting for the Euler equations. In *Proceedings of the 8<sup>th</sup> International Conference on Numerical Methods in Fluid Dynamics*, pages 507-512, Springer-Verlag, 1982.
  - 5 M. S. Liou and C. J. Steffen. A new flux splitting scheme. *Journal of Computational Physics*, 107:23-39, 1993.
  - 6 S. Lui and K. Xu. Entropy analysis of kinetic flux vector splitting schemes for the compressible Euler equations. ICASE Report No. 99-5, 1999.
  - 7 J. C. Mandal and K. Jain. A new implicit formulation of KFVS scheme for Euler equations. In *36<sup>th</sup> AIAA Fluid Dynamics Conference and Exhibit*, AIAA-2006-3709, 2006.
  - 8 S. K. Godunov. A finite difference method for the computation of discontinuous solutions of the equations of fluid dynamics. *Mat. Sb.* 47:357-393, 1959.
  - 9 P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*. 43:357-372, 1981.
  - 10 E. F. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock Waves*. 4:25-34, 1994.
  - 11 K. Xu. A Gas-Kinetic BGK Scheme for the Navier-Stokes equations and its connection with artificial dissipation and Godunov method. *Journal of Computational Physics*, 171:289-335, 2001.

## **The Cal Poly spacecraft attitude dynamics simulator**

Project Investigators:

Eric A. Mehiel and Carson O. Mittelsteadt  
Aerospace Engineering  
California Polytechnic State University  
San Luis Obispo, California

**The Cal Poly Spacecraft Attitude Dynamics Simulator (CP/SADS) is a four reaction wheel, pyramid configuration momentum exchange device which simulates the attitude dynamics and kinematics of a rigid body in space. The structure of the reaction wheel pyramid rests on a spherical air bearing. When the pivot point of the bearing is aligned with the center of gravity of the pyramid structure, no gravity torque is incumbent on the simulator. The combination of no gravitational torque on the structure and the four wheel momentum exchange device allows for dynamically realistic simulations of the attitude dynamics and kinematics of spacecraft in earth orbit. The goal of the CP/SADS project is to develop a test bed for spacecraft attitude control development, actuator and sensor design and characterization, inertial navigation sensor and algorithm development, and precision pointing algorithm development. This paper outlines the design components of the CP/SADS which are unique. Specifically, the CP/SADS uses a four axis pyramid configuration instead of the more typical orthogonal three axis configuration. The CP/SADS employs onboard power and will soon take advantage of wireless data transmission.**

**From the beginning of the project, performance requirements were internally developed and imposed. The CP/SADS project is also inherently multidisciplinary and students with expertise from departments outside the Cal Poly Aerospace Engineering department were utilized. However, the CP/SADS project was managed by students and faculty from the Aerospace Engineering department. The structure of the simulator was designed and built by Cal Poly Mechanical Engineering Students. The reaction wheels, all supporting electronics, and software required for operation, were designed and developed by Cal Poly Aerospace Engineering graduate students and Cal Poly faculty.**

**The simulator preliminary performance results are presented and discussed. Currently the simulator provides motion in only one axis since the final balancing of the system has not been performed. The technique used to balance the simulator in all three axes is still in development. Two tests were conducted on the simulator to show its performance capabilities. The first test demonstrated the capability of the simulator to maneuver from arbitrary initial position to some desired final position. The second test showed the ability of the simulator to maintain a fixed position when perturbed. The results of the platform correlated well with the predicted results found through analytical analysis and simulation. The test results were also used to verify the performance requirements of the CP/SADS were met, including a required maximum angular velocity of 0.1 rad/sec and a maximum acceleration of 0.1 rad/sec/sec.**

## **I. Spacecraft Attitude Dynamic Simulators and Momentum Exchange Devices**

The primary purpose of a satellite attitude control system is to stabilize and to provide attitude control for the spacecraft. Momentum exchange devices are one method commonly used to provide the control torques necessary for accurate attitude control. One type of momentum exchange device commonly employed is known as a Reaction Wheel. Reaction Wheels operate by exchanging an equal and opposite amount of angular momentum between the fly wheel of the Reaction Wheel and the system. Reaction Wheels exchange angular momentum with the body of the spacecraft by changing the angular velocity of the fly wheel. Hence, the direction of the angular velocity vector remains constant while the magnitude of the vector changes. When several Reaction Wheels are combined onto one spacecraft platform, the changes in the different wheel speeds are combined to affect the attitude of the spacecraft in all three axes. In general a fully agile spacecraft needs three reaction wheels at a minimum and sometime four are used to provide redundancy of the attitude control system.

Reaction Wheels differ from other Momentum Exchange devices such as Control Moment Gyroscopes, CMGs, in design but not in principle. Control Moment Gyros exchange momentum with the spacecraft by changing the direction of the angular velocity vector of the flywheel while keeping the direction constant. The dynamic analysis of Reaction Wheels is typically easier since the Inertia Matrix of the system does not vary with time since Reaction Wheels are typically designed to be axisymmetric. In contrast, CMGs momentum exchange devices usually cause the inertia matrix of the system to be time varying. There are also several groups researching Variable Speed Control Moment Gyros which combine the gyroscopic effect of classical CMGs and the angular velocity effect of Reaction Wheels. Regardless of the momentum exchange device employed, the stability analysis of both Reaction Wheels and CMGs requires nonlinear stability theory when a fully agile system is desired. Both Reaction Wheels and CMGs suffer from saturation and gimbal lock problems which make controller design difficult at best.

While the dynamic analysis of momentum exchange devices is a challenge in itself, designing control laws for Reaction Wheels and CMGs is made that much more difficult due to the often time varying and nonlinear nature of the equations of motion. Therefore, it has been historically common to linearize the equations of motion around a single operating point, or to command the spacecraft in such a way as to decouple the angular velocities about orthogonal axes. Recently several groups have developed full state feedback laws which guarantee asymptotic stability. For example, Wie has shown asymptotic stability can be achieved with full state feedback when different conditions on the inertia matrix of the system are satisfied. To overcome gimbal lock and reaction wheel speed saturation, several approaches have been developed. See for example...

The adage, "a proof is worth a thousand simulations", remains very true in the world of control theory and controller design. If a stability proof can be formulated, all dynamic simulations should experimentally demonstrate stability properties. However, to gain confidence in a particular control law in a real world setting complete with noise, other disturbances and errors associated with measurements and control authority, groups typically abide by the accepted practice of building numeric and physical simulations.

The Cal Poly Spacecraft Dynamic Simulator, CP/SADS, has been designed to serve as a way of quickly assembling experimental data while following the three pronged approach outlined above. Ideally, theoretical analysis would precede numeric simulation followed by experimental data and finally, the spacecraft could be built. This paper outlines the design and development of

such an experimental test bed. The kinematic and dynamic analysis used to predict the motion of the simulator is presented. There are significant differences in the dynamics of rigid bodies in space and rigid bodies on ground. The constant freefall of spacecraft in earth orbit is the most significant difference. Spacecraft simulators on the surface of the earth must be supported and hence a gravity torque is common. Our approach to this significant problem will be discussed in more detail. The design and development of the CP/SADS will also be presented. Each subsystem will be covered in detail, including all structural components, the reaction wheels, power supply, sensors and actuators, data acquisition, and real time control. The preliminary results and testing procedures developed to date will also be presented. While the CP/SADS is still under development, preliminary test results agree well with predicted performance. Finally, the future of the CP/SADS will be outlined. Our group hopes to develop a 'plug-and-play' approach based on using the CP/SADS for attitude control algorithm testing.

#### **A. Past and Current Spacecraft Dynamic Simulators**

Reaction wheel control devices have been around since the Apollo missions in the 1960's. The four-wheel pyramidal reaction wheel configuration is a common type of reaction wheel system and is used in many spacecraft today. The advantages of a four-wheel pyramidal configuration for attitude control devices are its light weight and redundancy since the pyramidal configuration has four motors rather than the three that are minimally needed for full attitude control.

The Control Momentum Gyro (CMG) is a momentum exchange device that is closely related to the reaction wheel. Like the reaction wheel, the CMG was developed to make precise changes in attitude control of the spacecraft. The CMG differs from a reaction wheel in that the inertia wheel's spin is not fixed on a precise angle relative to the spacecraft. Instead the CMG's are nutated by a servomotor. The inertia wheels are rotated at constant angular velocity and the nutation creates a moment perpendicular to the spin and the nutation axes. CMG's have an advantage over Reaction Wheels since they are not limited by saturation of the motors. However, depending on the configuration CMG's can be plagued by gimbal lock, which occurs when two or more gimbal axes have been aligned. Reaction Wheels do not suffer from gimbal lock. Gimbal lock is usually combated by using multiple CMG's. Future versions of the Cal Poly Reaction Wheel Platform may incorporate CMG wheels.

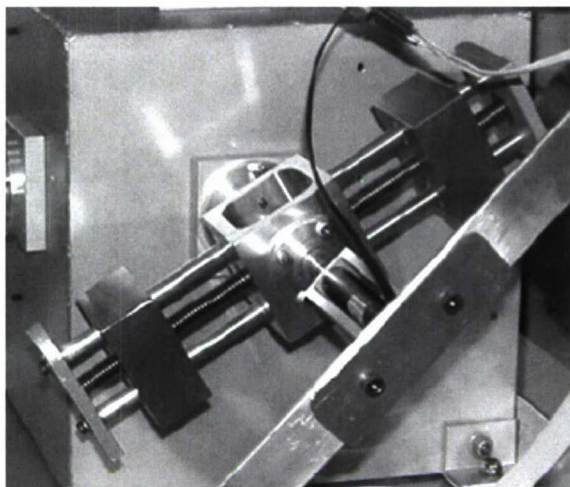
New and different reaction wheel control algorithms and products have been developed in recent years. A recent variation of the reaction wheel is the Variable Inertia Reaction Wheel (VIRW) developed by researchers at the Georgia Institute of Technology (Figure 1<sup>2</sup>) as part of the Variable Inertia Test Platform (VITP). The VIRW has the ability to change the inertias of any of the wheels. It does this by way of two masses which are allowed to transverse the length of the wheel while it is spinning. This type of system requires balancing techniques that are complicated and difficult to implement. However, if the added balancing complexity is tolerable, the VIRW provides two distinct advantages: variable precision and vehicle reorientation without the change in angular speed of the wheel. Variable precision is critical in reducing jitter, which allows for fine pointing of the spacecraft.<sup>2</sup>



a) Complete VITP

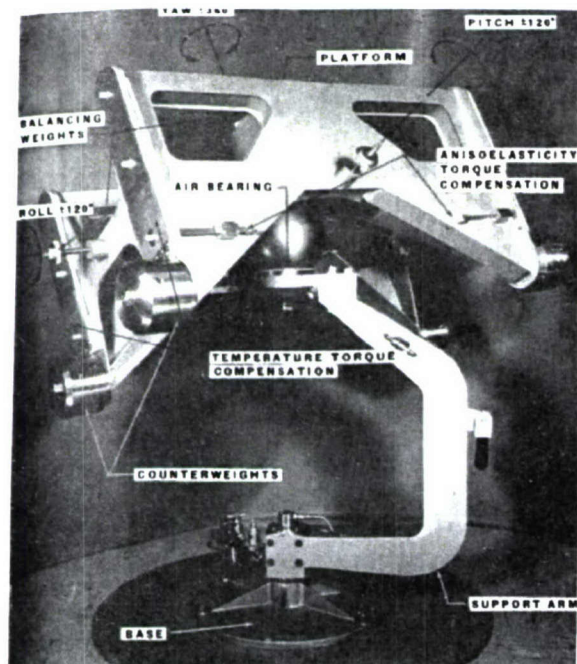


b) VITP with foam corner guards removed



**Figure 1: Images of variable inertia test platform**

Regardless of what type of momentum exchange device is developed, to simulate the rotational motion associated with spacecraft in orbit, the momentum exchange device is typically mounted on some sort of bearing surface. The combination of a reaction wheel of any flavor or a CMG with two or three axis low friction bearing surface allows for spacecraft attitude dynamics simulations. Figure 2 is a photo of one of the first spacecraft dynamics simulator. The Satellite Motion Simulator developed by NASA in the 1960's was used as a test bed to design and develop satellite attitude control laws during the early days of the United States Space Program.

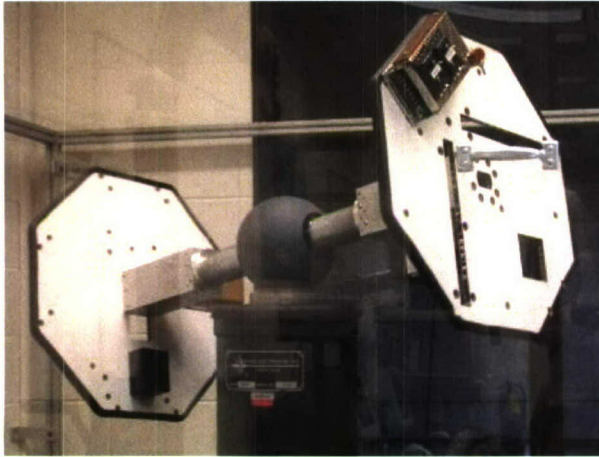


**SATELLITE MOTION SIMULATOR**

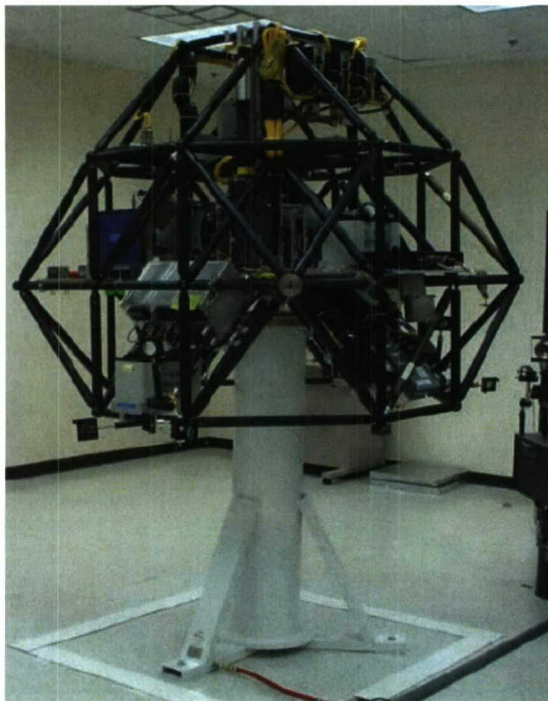
**Figure 2: 1960 NASA Air Bearing Satellite Simulator**

Figure 3 is a recent photo taken from the Virginia Tech Space Systems Simulation Laboratory (SSSL) website<sup>10</sup>. The SSSL maintains several space dynamics simulators, the Distributed

Spacecraft Attitude Control System Simulator (DSACS), Whorl-I and Whorl-II, pictured in Figure 3. DSACS uses a spherical air bearing to simulate the attitude dynamics of a spacecraft while Whorl-I and Whorl-II employ a dumbbell air bearing configuration to achieve 360 degrees of rotational motion in two axes. All three systems are used to simulate the attitude dynamics of spacecraft in orbit. Finally, Figure 4, is a photo of the Honeywell Space System Simulator currently under development. Notice the Honeywell simulator also has a modular design.



**Figure 3: Virginia Tech Dumbbell Style Simulator**

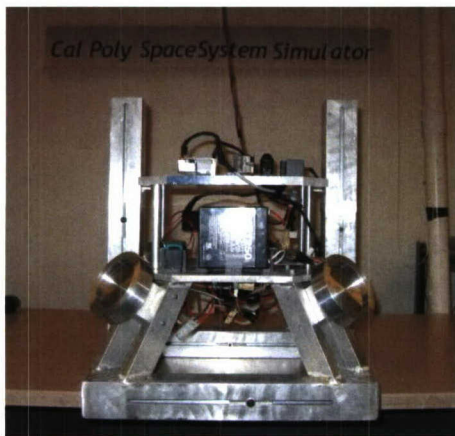


**Figure 4: Honeywell Space System Simulator, circa 2003**

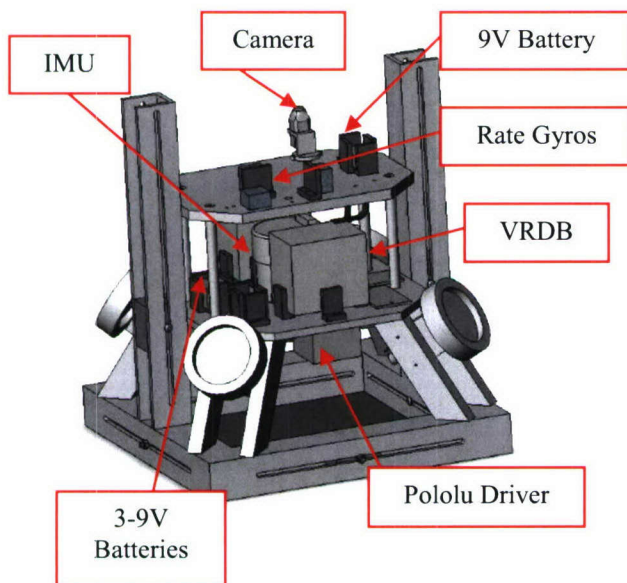
### **B. An Overview of the Cal Poly Spacecraft Attitude Dynamics Simulator**

A recent photo of the CP/SADS is shown in Figure 5. A solid model of the CP/SADS is shown in Figure 6. Notice the CP/SADS structure rests on a spherical air bearing. Hence the CP/SADS has a full 360 degrees of motion about the z-axis and  $\pm 30$  degrees about both the x-

and y-axes. The CP/SADS differs from the projects discussed above since it relies on a pyramidal reaction wheel configuration. The reaction wheels are placed along the edges of what amounts to a square pyramid. The pyramidal configuration of the reaction wheels and the associated structure is inherently symmetric about all three axes. Only the relatively small masses associated with electronic components and the power supply have to be balanced about the center of gravity of the air bearing. More discussion on balancing is found below. The CP/SADS also has a coarse balancing system built into the structure which is used to modify the center of gravity of the CP/SADS. Appendages can be added to the platform to test different satellite configurations such as solar arrays. More detail on all of the CP/SADS subsystems and design can be found in subsequent sections.



**Figure 5: The Cal Poly Spacecraft Attitude Dynamics Simulator**



**Figure 6: Solid Model of the Cal Poly Spacecraft Attitude Dynamics Simulator**

## II. Spacecraft Equations of Motion: In Space vs. the Simulator

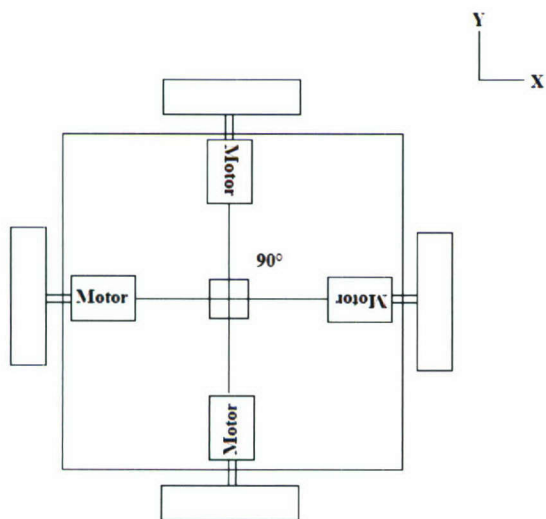
### A. Dynamics

In this section the equations of motion of a satellite are developed. The Dynamic equations and the kinematic equations used to model the CP/SADS are presented. The significant differences in the equations of motion of a spacecraft in orbit and a spacecraft simulator on the ground are discussed in different sections. The approach taken here to model the dynamics of the simulator follows a fairly standard derivation of Euler's Equations for rigid body rotational dynamics. To model the kinematics of the simulator, a quaternion representation is developed.

#### 1. Reaction Wheel Dynamics

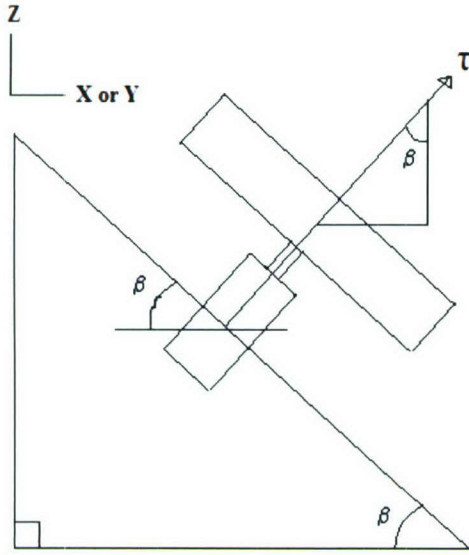
The pyramidal reaction wheel design is favored over a three-axis reaction wheel design due to its inherent symmetry and increase in torque along each rotational axis. The symmetry and increased torque come with the price of added mass. Therefore, a final design will generally depend on the nature of the problem and any constraints applied to the system. A four axis reaction wheel design also adds redundancy in case one motor fails before the mission is complete.

When looking at the x and y axes as shown in Figure 7, all four wheels can be arranged at equal angles from each other ( $90^\circ$  apart) which allows for equal torque in the x and y planes. Since the platform is approximately symmetric in the x and y axes the principal moments of



**Figure 7: The four motors arranged  $90^\circ$  apart on the x and y axes**

inertia in the x and y directions are assumed equal.



**Figure 8: Motor tilted into the z axis by angle,  $\beta$**

To apply a moment in only the z-axis all four, or any opposite two motors, are actuated simultaneously. The motor tilt angle  $\beta$  is formed as the angle from the motor spin axis to the z-axis. If the proper motor tilt angle is used, then the moment applied in the z-direction when all four motors accelerate is equal to the moment applied when any opposite two motors apply a moment about the axis connecting the two motors. Hence, due to symmetry, with the right choice of motor tilt angle, the moments applied in all three axes are equal.

Since the reaction wheel motors are tilted up toward the z-axis, all four motors contribute to the moment applied along the z-axis while only two motors are capable of applying a torque in either the x- or y- direction. When the applied moment from each motor is accounted for, the total moment in the x-, y-, and z-directions is expressed as,

$$\begin{aligned} T_x &= J(\dot{\omega}_{w_1} + \dot{\omega}_{w_3})\sin(\beta) = I_x \dot{\omega}_x \\ T_y &= J(\dot{\omega}_{w_2} + \dot{\omega}_{w_4})\sin(\beta) = I_y \dot{\omega}_y \\ T_z &= J(\dot{\omega}_{w_1} + \dot{\omega}_{w_2} + \dot{\omega}_{w_3} + \dot{\omega}_{w_4})\cos(\beta) = I_z \dot{\omega}_z \end{aligned} \quad (1)$$

where the inertia of each wheel is the same and given as,  $J$ . The angular acceleration of each wheel is given by  $\dot{\omega}_{w_j}$  and the angular acceleration of the platform in the x-, y- and z-axes are given by  $\dot{\omega}_x$ ,  $\dot{\omega}_y$  and  $\dot{\omega}_z$  respectively. The principal moments of inertia of the platform, including the inertia of each wheel about the platform center of gravity, are given by  $I_x$ ,  $I_y$  and  $I_z$ . The principal axes lie close to the body frame of reference demonstrated in Figures 7 and 8 due to the symmetry of the platform.

If the angular acceleration in the z-axis is equal to the angular accelerations in either the y or x axes then the reaction wheel set has equal control authority in all three axes. Assuming the angular acceleration is the same for each reaction wheel and setting the angular acceleration in the x- or y-direction equal to the angular acceleration in the z-direction as given in Eq. (1) leads to the relationship,

$$\frac{\sin(\beta)}{I_{x,y}} = \frac{2\cos(\beta)}{I_z}. \quad (2)$$

Solving for the angle  $\alpha$  which insures equal control authority results in,

$$\alpha = \tan^{-1}\left(2\frac{I_{x,y}}{I_z}\right). \quad (3)$$

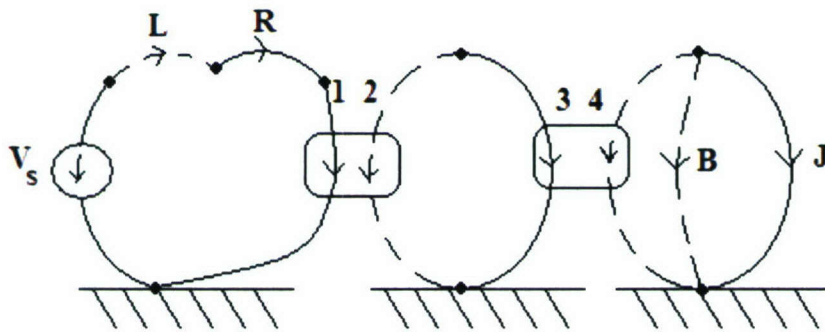
The ideal motor tilt angle is dependent on the relationship between the x- and y- principal inertias and the z- principal inertia. If  $I_{x,y} = I_z$  the angle,  $\beta$ , is  $63.4^\circ$ . However, if

$I_{x,y} < I_z$  then  $\beta$  is less and conversely if  $I_{x,y} > I_z$  then  $\beta$  is greater.

Initial design concepts of the CP/SADS indicated that the inertia about the z-axis would be greater than the inertias in the x- and y-axes. The inertias in the x- and y-axes were assumed equal. Therefore, a motor tilt angle of  $57.0^\circ$  was chosen for  $\alpha$  and written down as a hard requirement in the specifications list given to the structures team<sup>5</sup>. The final design advocated by the structures team had an inertia in the z-direction of  $0.5952 \text{ kg-m}^2$  and  $0.5528 \text{ kg-m}^2$  in the x and y directions. Hence, the ideal motor tilt angle is actually  $61.7^\circ$ . Since the ideal motor tilt angle is slightly greater than  $57.0^\circ$  the CP/SADS has slightly more control authority in the z-axis than in the x- and y-axes.

## 2. Actuators Dynamics

In order to achieve an angular acceleration of  $0.1 \text{ rad/sec}^2$  and an angular velocity of  $0.1 \text{ rad/sec}$  as given in the specifications list a detailed analysis was done to insure proper motor sizing. Using linear graphing methodology a state equation was developed to relate input voltage to output torque on the platform<sup>3</sup>. Figure 9 shows the linear graph of a motor with a gear train.



**Figure 9: Linear graph of motor with gear train**

The moment of inertia,  $J$ , is the combination of the reaction wheel inertia and motor shaft inertia. Viscous damping is given by  $B$ . The resistance of the motor windings is represented by  $R$  and inductance produced by the current flowing around the winding is represented by  $L$ . The input voltage is given by  $V_s$ . The two transducers shown by the numbers 3-4 and 1-2 are the two-port elements which represent the gear train and the motor respectively. Since the

transduction process relates across-variables to across-variables the connection between both 1-2 and 3-4 are related by two-port transformers defined by,

$$\begin{bmatrix} v_1 \\ f_1 \end{bmatrix} = \begin{bmatrix} TF & 0 \\ 0 & -1/TF \end{bmatrix} \begin{bmatrix} v_2 \\ f_2 \end{bmatrix} \quad (4)$$

Equation (4) relates the through-variables,  $f$ , to the across-variables,  $v$ . The transformer ratio is given by,  $TF$ , where  $TF$  equals the back-EMF constant,  $K_E$ , and the torque constant,  $K_M$ , for the motor. The two-port transformer relationship for the motor is then given as,

$$\begin{bmatrix} V_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} K_E & 0 \\ 0 & -1/K_M \end{bmatrix} \begin{bmatrix} \Omega_2 \\ T_2 \end{bmatrix}. \quad (5)$$

The transformer ratio also equals the gear ratio,  $GR$ , for the gear train. The two-port transformer relationship for the gear train is also given as,

$$\begin{bmatrix} \Omega_3 \\ T_3 \end{bmatrix} = \begin{bmatrix} GR & 0 \\ 0 & -1/GR \end{bmatrix} \begin{bmatrix} \Omega_4 \\ T_4 \end{bmatrix}. \quad (6)$$

From the linear graph, eight elemental equations describe the energy flow and are given as,

$$\begin{aligned} \frac{di}{dt} &= \frac{1}{L} V_L, & \frac{d\Omega_J}{dt} &= \frac{1}{J} T_J, \\ T_B &= B\Omega_B, & V_R &= Ri_R, \\ V_1 &= K_E \Omega_2, & T_2 &= -K_M i_1, \\ \Omega_3 &= GR\Omega_4, & T_4 &= -GRT_3, \end{aligned} \quad (7)$$

Along with the elemental equations we can write down five compatibility equations given as,

$$\begin{aligned} V_1 &= -V_R - V_L + V_S \\ \Omega_2 &= \Omega_3 \\ \Omega_4 &= \Omega_B \\ \Omega_B &= \Omega_J \\ \Omega_4 &= \Omega_J \end{aligned} \quad (8)$$

From the linear graph we also get five continuity equations given as,

$$\begin{aligned}
i_1 &= i_R \\
i_L &= i_R \\
i_2 &= i_3 \\
T_J &= -T_4 - T_B \\
T_3 &= -T_2
\end{aligned} \tag{9}$$

The combination of the elemental equations (7), compatibility equations (8) and the continuity equations (9) results in the two coupled differential equations,

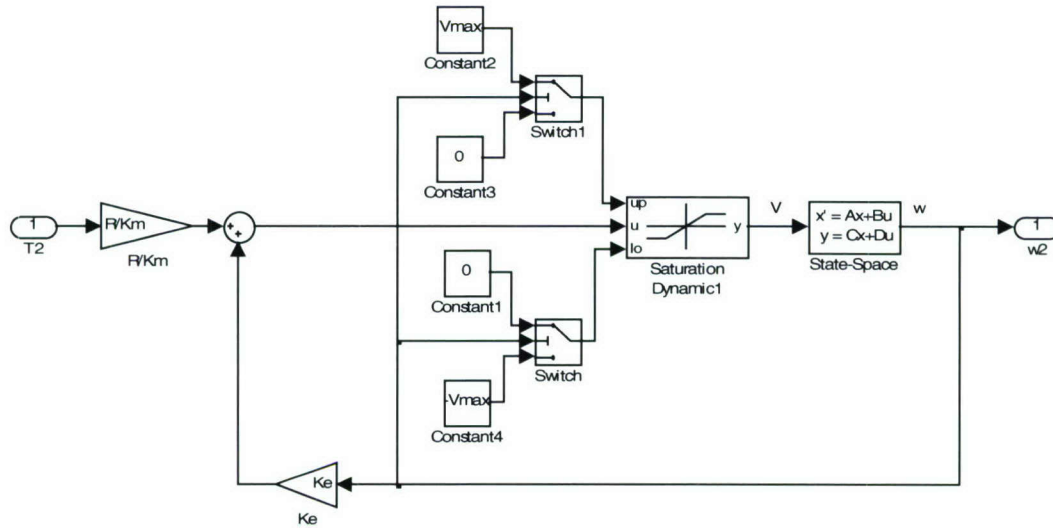
$$\begin{aligned}
\frac{di_L}{dt} &= \frac{1}{L}(-K_E \cdot GR \cdot \Omega_J - R \cdot i_L + V_S) \\
\frac{d\Omega_J}{dt} &= \frac{1}{J}(K_M \cdot GR \cdot i_L - B \cdot \Omega_J)
\end{aligned} \tag{10}$$

The two first order ordinary differential equations from Eq. (10) are written in state space form as,

$$\frac{d}{dt} \begin{bmatrix} i_L \\ \Omega_J \end{bmatrix} = \begin{bmatrix} \frac{-R}{L} & -GR \frac{K_E}{L} \\ GR \frac{K_M}{J} & \frac{-B}{J} \end{bmatrix} \begin{bmatrix} i_L \\ \Omega_J \end{bmatrix} + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} V_S \tag{11}$$

The angular velocity of each reaction wheel is found by integrating Eq. (11). Finally, the angular momentum of each reaction wheel is determined by multiplying the moment of inertia of each reaction wheel with the appropriate angular velocity. Since angular momentum is conserved between the reaction wheel and the platform the angular velocity of the platform is calculated after taking into account the number of reaction wheels and their orientation with respect to the platform's body frame.

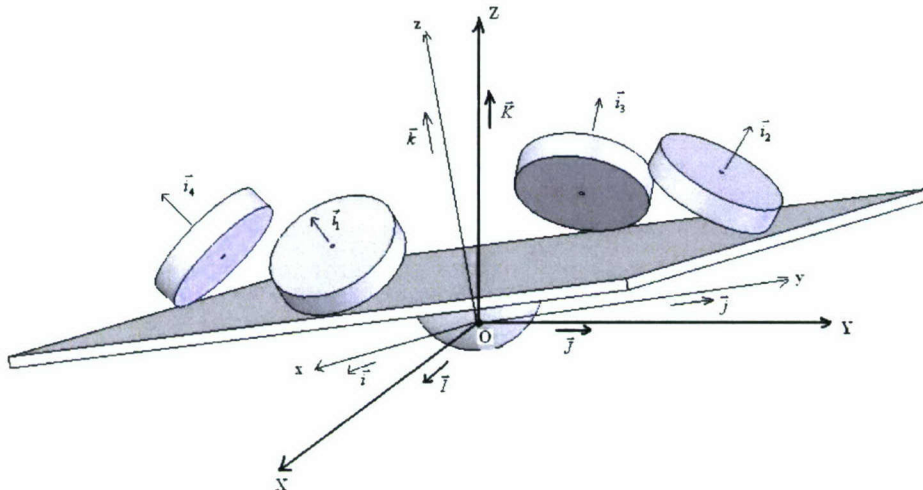
Figure 10 is a MatLab-Simulink block diagram which takes the input torque command and determines the angular velocity of each the reaction wheel. When a torque is commanded to the motor the code takes the torque and converts this torque into an equivalent input voltage will drive the motor to the commanded torque. Each motor accepts a voltage ranges from 0V to +24V when the motor is spinning in the positive direction and 0V to -24V when the motor is spinning in the negative direction. When the motor is not spinning the voltage range is from +24V to -24V. The two trigger switches change the dynamic saturation minimum and maximum depending on the speed of the motor. The voltage applied to the motor is not allowed to change polarity when the motor is spinning since such an applied voltage would create a large current spike and damage the insulation on the wire in the motor coil. The state-space block in Figure 5 represents Eq. (11) and transfers the voltage to an angular velocity of the reaction wheel.



**Figure 10: Motor simulation with reaction wheel attached**

### 3. CP/SADS Dynamics

To derive the dynamic equations of motion of the CP/SADS, two assumptions were made. First, the center of gravity of the CP/SADS structure was aligned with the pivot point of the spherical air bearing supporting the structure. Therefore, the gravity torque due to misalignments is zero. Second, the coordinate system shown in Figure 11 (adapted from Healy<sup>4</sup>) with basis vectors  $\vec{I}, \vec{J}, \vec{K}$  is an inertial reference frame. Next, the body fixed frame, B, was defined by the basis vectors  $\vec{i}, \vec{j}, \vec{k}$ . Both frames have fixed origins at the center of rotation of the air bearing, O. The unit vectors through the axis of rotation of the reaction wheels are  $\{\vec{i}_1, \vec{i}_2, \vec{i}_3, \vec{i}_4\}$ .



**Figure 11: Reference Frame Schematic**

Using the inertial and body frames the equations of a motion for a rigid body are derived from Euler's Equation where the applied torque must equal the time rate of change of the angular

momentum in the inertial frame. The angular momentum of a rigid body denoted by  $\vec{h}$  is given as,

$$\vec{h} = I \cdot \vec{\omega} \quad (12)$$

In Eq. (12), the platform's inertia matrix,  $I$ , is not diagonal; it includes the products of inertia as well as the moments of inertia and is given by,

$$I = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy} & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz} \end{bmatrix}. \quad (13)$$

In the case of the CP/SADS the rigid body has four reaction wheels, therefore, Eq. (12) must be expanded to include the angular momentum of each reaction wheel. Hence,

$$\vec{h} = I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)}. \quad (14)$$

Equation (14) states that the angular momentum  $\vec{h}$  equals the product of the inertia matrix of the entire platform about the platform center of gravity,  $I_G$ , with the angular velocity of the platform,  $\vec{\omega}^{(p)}$ , added to the sum of each reaction wheel's angular momentum about the motor spin axis,  $J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)}$ .

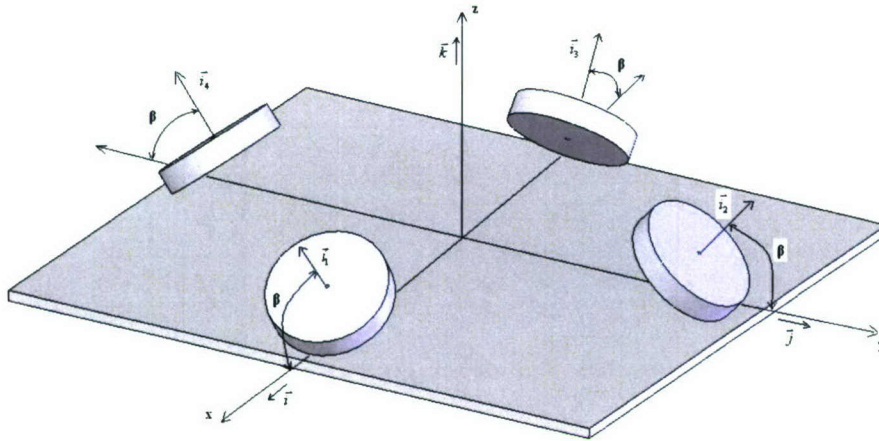
Taking the time derivative of the angular momentum in Eq. (14) gives the moment equation with respect to the inertial frame. By applying the proper derivative and cross product to account for the rotating body reference frame used to express the inertias, the dynamic equations of motion of the CP/SADS is expressed as,

$$\vec{M} = \frac{d}{dt} \left( I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)} \right) + \vec{\omega}^{(p)} \times \left( I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)} \right). \quad (15)$$

Since all four reaction wheels on the platform are identical the moment of inertia of each wheel given by  $J_{G_n}^{(n)}$  can be written as  $J$ . Rewriting Eq. (15) to separate out the wheel dynamics from the platform dynamics yields,

$$\vec{M} = I_G \cdot \dot{\vec{\omega}}^{(p)} + \vec{\omega}^{(p)} \times I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 \left( J \cdot \dot{\vec{\omega}}_{rel}^{(n)} + \vec{\omega}^{(p)} \times J \cdot \vec{\omega}_{rel}^{(n)} \right). \quad (16)$$

The summation in Eq. (16) is valid when the reaction wheel and platform angular velocity vectors are expressed in the same coordinate system. Therefore, a coordinate transformation is performed to rotate the angular velocity of each reaction wheel expressed in the wheel frame to the body frame. Figure 12 shows the orientation of each wheel with respect to the platform<sup>4</sup>.



**Figure 12: Reaction Wheel Pyramid Formation**

In Figure 12, the motor tilt angle is again represented by  $\beta$  and is the angle between each reaction wheel's spin axis and the corresponding body axis it lies on. To determine the angular velocity of each reaction wheel in the body frame, the following rotation matrices were developed to rotate from the reaction wheel spin axis to the body axis. The subscript  $n$  denotes the number of each wheel shown in Figure 12<sup>4</sup>.

$$\begin{aligned}
 R_{W1} &= \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_{W2} &= \begin{bmatrix} 0 & 1 & 0 \\ \cos \beta & 0 & -\sin \beta \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_{W3} &= \begin{bmatrix} -\cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_{W4} &= \begin{bmatrix} 0 & 1 & 0 \\ -\cos \beta & 0 & \sin \beta \\ \sin \beta & 0 & \cos \beta \end{bmatrix}
 \end{aligned} \quad (17)$$

Since the reaction wheels are symmetric their inertia matrix is diagonal in the reaction wheel axis. The major principal axis is the same as the torque axis. The angular momentum equation for each wheel in its own coordinate frame is given by,

$$h_n = J \cdot \vec{\omega}_{rel}^{(n)} = \begin{bmatrix} J_{11} & 0 & 0 \\ 0 & J_{22} & 0 \\ 0 & 0 & J_{33} \end{bmatrix} \begin{bmatrix} \omega_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{11} \omega_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} \quad (18)$$

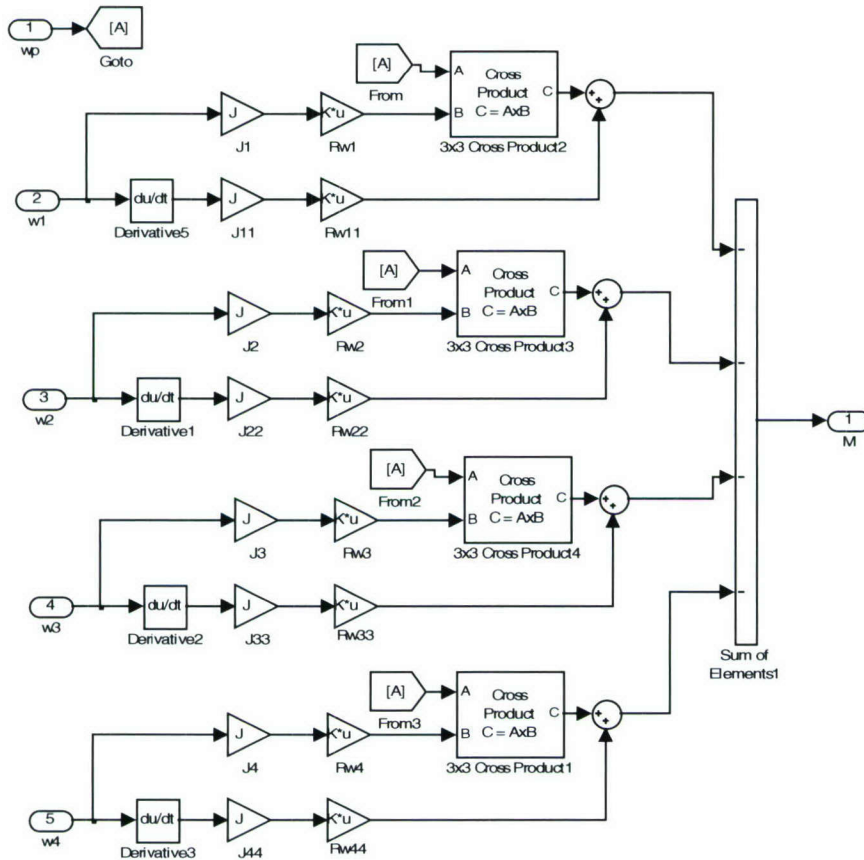
Taking the time derivative of the momentum equation gives the torque contribution of each reaction wheel.

$$T_n = J \cdot \dot{\vec{\omega}}_{rel}^{(n)} = \begin{bmatrix} J_{11} & 0 & 0 \\ 0 & J_{22} & 0 \\ 0 & 0 & J_{33} \end{bmatrix} \begin{bmatrix} \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{11} \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix}. \quad (19)$$

Finally, the equations of motion for the CP/SADS are found after substituting Eqs. (18) and (19), along with the coordinate rotations from Eq. (17) into Eq.(16).

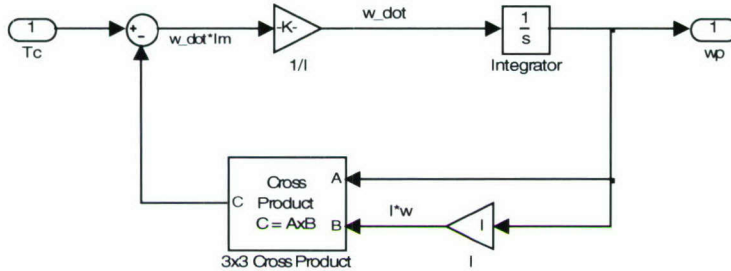
$$\vec{M} = I_G \dot{\vec{\omega}}^{(p)} + \vec{\omega}^{(p)} \times I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 \left( R_{W_n} \cdot \begin{bmatrix} J_{11} \cdot \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} + \vec{\omega}^{(p)} \times R_{W_n} \cdot \begin{bmatrix} J_{11} \cdot \omega_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} \right). \quad (20)$$

Figure 13 is a Simulink block diagram which represents the reaction wheel dynamics given in the summation term on the right hand side of Eq. (20). The output of the subsystem represented in Figure 13 is the sum of moments for all four wheels with respect to the inertial frame.



**Figure 13: Reaction wheel dynamic simulation**

Figure 14 is the Simulink block diagram which models the dynamics of the platform and corresponds to the terms outside the summation on the right hand side of Eq. (20). In general the Simulink model allows for commanded torques from the reaction wheels and any external moment.



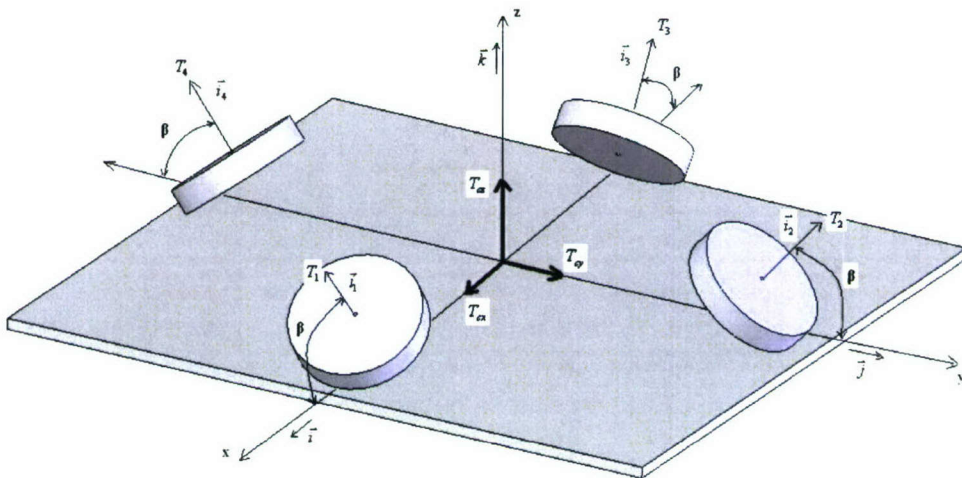
**Figure 14: Dynamic platform simulation**

#### 4. Commanded Torque Distribution and Minimization

Since the CP/SADS has four reaction wheels the commanded torque requirement should be distributed to each motor in such a way to minimize the energy used by the motors. With respect to the geometry of the platform in shown in Figure 15 the command torque on each body axis is,

$$\begin{aligned} T_{cx} &= T_1 \cos(\beta) - T_3 \cos(\beta) \\ T_{cy} &= T_2 \cos(\beta) - T_4 \cos(\beta) \\ T_{cz} &= T_1 \sin(\beta) + T_2 \sin(\beta) + T_3 \sin(\beta) + T_4 \sin(\beta) \end{aligned} \quad (21)$$

The torques,  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$  are the torques about each reaction wheel spin axis.



**Figure 15: Torque Relationships on a Reaction Wheel Pyramid**

In matrix form Eq. (21) is represented as,

$$\begin{bmatrix} \hat{T}_{cx} \\ \hat{T}_{cy} \\ \hat{T}_{cz} \end{bmatrix} = \begin{bmatrix} T_{cx} / \cos(\beta) \\ T_{cy} / \cos(\beta) \\ T_{cz} / \sin(\beta) \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = [A_w]T \quad (22)$$

Since the matrix  $A_w$  is not square it cannot be inverted. In order to find the motor torques of each motor an auxiliary equation is added. To minimize the required torque each motor applies, we can use the method of Calculus of Variations to minimize the norm of the torque vector,  $\vec{T}$ . Minimization of the norm of the vector  $\vec{T}$  is accomplished using the Lagrangian,

$$L = H + \lambda_1 g_1 + \lambda_2 g_2 + \lambda_3 g_3. \quad (23)$$

The Hamiltonian,  $H$ , from Eq. (23) is expressed as,

$$H = \sum_{n=1}^4 T_n^2 \quad (24)$$

Therefore, the Lagrange coefficients from Eq. (23) are derived from Eq. (22) and Eq. (24) as,

$$\begin{aligned} g_1 &= T_1 - T_3 - \hat{T}_{cx} \\ g_2 &= T_2 - T_4 - \hat{T}_{cy} \\ g_3 &= T_1 + T_2 + T_3 + T_4 - \hat{T}_{cz} \end{aligned}$$

The partial derivatives of the Lagrangian with respect to the torque on each motor generate the solution which satisfies the conditions for minimizing the Hamiltonian. The partial derivatives are,

$$\begin{aligned} \frac{\partial L}{\partial T_1} &= 2T_1 + \lambda_1 + \lambda_3 = 0 \\ \frac{\partial L}{\partial T_2} &= 2T_2 + \lambda_2 + \lambda_3 = 0 \\ \frac{\partial L}{\partial T_3} &= 2T_3 - \lambda_1 + \lambda_3 = 0 \\ \frac{\partial L}{\partial T_4} &= 2T_4 - \lambda_2 + \lambda_3 = 0 \end{aligned} \quad (25)$$

The solution to Eq. (25) gives the desired auxiliary equation so the state matrix from Eq. (22) can be inverted. The final auxiliary condition is,

$$T_1 - T_2 + T_3 - T_4 = 0.$$

Therefore, Eq. (22) is rewritten in matrix form and inverted to give the relationship between the desired torque command to the rigid body and the torque command to each reaction wheel motor. The final reaction wheel torque command is given as,

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \hat{T}_{cx} \\ \hat{T}_{cy} \\ \hat{T}_{cz} \\ 0 \end{bmatrix}.$$

With the command torque for each reaction wheel determined, the dynamics of the CP/SADS are determined, if the simulator were in space. However, there are significant torques present due to gravity which must be considered when the simulator is in the lab, on earth.

### 5. Ground Simulation

The derivation of the dynamics of the CP/SADS in the above section assumes the pivot point of the spherical air bearing holding the CP/SADS is collocated with the center of gravity of the CP/SADS. The CP/SADS is balanced when the center of gravity and the pivot point of the CP/SADS are collocated. While balancing the CP/SADS is a desired goal, the location of the center of gravity is not easy to find. Relying on System Identification techniques, and following the methods developed by Tanygin and Williams<sup>11</sup>, Healy<sup>4</sup> has shown that the center of gravity of the CP/SADS can be located, and hence, moved by some sort of balancing technique. However, to date, the System Identification technique has not been applied to the CP/SADS. The development of the balancing device and algorithm is ongoing.

To compensate for the unbalanced simulator the preliminary results discussed in subsequent sections only considers rotations about the z-axis of the CP/SADS simulator. The symmetric design of the SP/SADS guarantees that the CP/SADS is nearly balanced in the x- and y-axes. Therefore, the coarse balancing system, (discussed below), was used to move the center of gravity of the CP/SADS well below the pivot point of the air bearing. Hence, the CP/SADS was passively stable about the x- and y-axes. Essentially, the CP/SADS behaves much like a pendulum about the x- and y-axes.

For a full analysis and discussion of the gravity torque introduced into the equations of motion of the simulator due to the dislocation of the center of gravity and the pivot point, see the paper by Healy<sup>4</sup>.

### B. Kinematics: Quaternions

To represent the kinematics of the CP/SADS a quaternion representation was used. The quaternion representation does not suffer from singularities in the solution to the kinematic equations. Also, the quaternion representation yields to nonlinear full state feedback stability analysis as described by Wie<sup>1</sup>. If the Euler angles are required for further analysis, the quaternion representation can yield that information.

In his book, Wie shows that the angular velocities, or body rates, of a rigid body in space are related to the quaternion and the quaternion derivatives by,

$$\begin{aligned}
\omega_1 &= 2(\dot{q}_1 q_4 + \dot{q}_2 q_3 - \dot{q}_3 q_2 - \dot{q}_4 q_1) \\
\omega_2 &= 2(\dot{q}_2 q_4 + \dot{q}_3 q_1 - \dot{q}_1 q_3 - \dot{q}_4 q_2) \\
\omega_3 &= 2(\dot{q}_3 q_4 + \dot{q}_1 q_2 - \dot{q}_2 q_1 - \dot{q}_4 q_3)
\end{aligned} \quad (26)$$

Since a quaternion is only a rotation when it's norm is equal to one, the derivative of the norm must then equal zero and is given as,

$$0 = 2(\dot{q}_1 q_1 + \dot{q}_2 q_2 + \dot{q}_3 q_3 + \dot{q}_4 q_4). \quad (27)$$

Combining Eqs. (26) and (27) into a vector matrix product gives,

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} q_4 & q_3 & -q_2 & -q_1 \\ -q_3 & q_4 & q_1 & -q_2 \\ q_2 & -q_1 & q_4 & -q_3 \\ q_1 & q_2 & q_3 & q_4 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \\ \dot{q}_4 \end{bmatrix}. \quad (28)$$

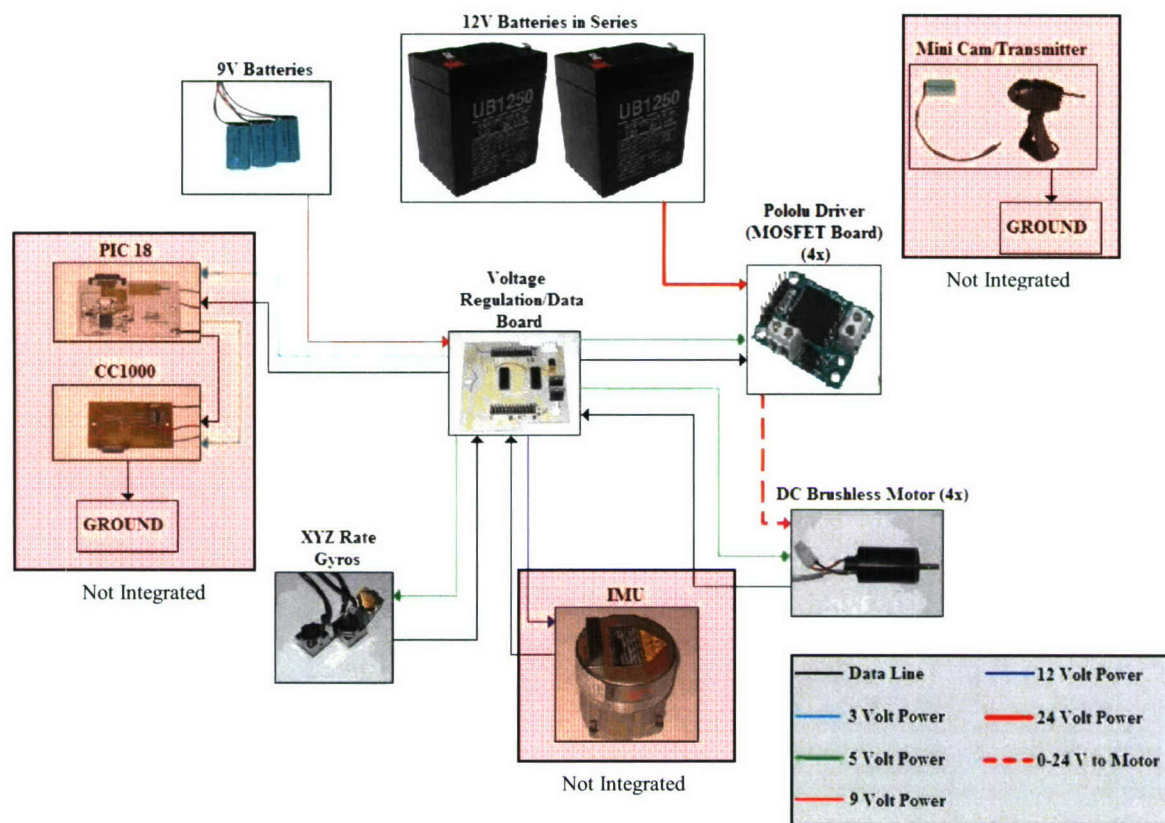
Taking the inverse of the quaternion matrix in order to relate the quaternion rates to the angular velocities and separation the vector part of the quaternion from the scalar part yields the differential equations,

$$\begin{aligned}
\dot{\vec{q}} &= \frac{1}{2}(\vec{q}_4 \vec{\omega}_p - \vec{\omega}_p \times \vec{q}) \\
\dot{q}_4 &= -\frac{1}{2} \vec{\omega}_p^T \vec{q}
\end{aligned} \quad (29)$$

### III. The Cal Poly Spacecraft Attitude Dynamic Simulator Design

#### A. Overall System Design

The CP/SADS is composed of five different sub-systems which include: the structure of the platform, the reaction wheel actuators, the power buses, sensors, and the data acquisition system. Figure 16 shows the entire hardware schematic of the platform. The voltage regulation/data board (VRDB) acts as the main hub for all low power and data lines. A 9 V battery is used to power the camera and its transmitter. The 12 V bus shown in Figure 16 is for future reference once the LN-200 IMU (Inertia Measurement Unit) is operational. The LN-200 was kindly donated by Northrop Grumman. The components behind the transparent boxes are components which will be added to the CP/SADS during future development stages.



**Figure 16: CP/SADS hardware schematic**

## B. Structural Design

The structure of the CP/SADS is used to hold and secure each sub-system; it is the body of the simulated spacecraft. Much of the design of the structure was determined by a specifications list. The specifications list was developed for the purpose of insuring hard requirements to the CP/SADS. Since most of the manufacturing and design was conducted by a mechanical engineering senior project team (see Martin, Quinto, and Nash<sup>5</sup>) a requirements document was established for clear communication between the Structure group and the System Integration group.

Many requirements were contemplated during the development of the specifications list; however, only five important requirements of the platform will be mentioned in this report. Important requirement #1 limited the weight to less than 10 Kg without including the adjusting weights, (discussed below). This requirement allowed for one user to pick up and operate the platform. The structure essentially acts as a simulated spacecraft structure and allows for mounting of other sensors, actuators, and other subsystems.

The structure was designed with angled supports which hold the reaction wheel motors at  $57^\circ$ . With the supports angled at  $57^\circ$  the reaction wheels would be angled at  $33^\circ$  from horizontal giving the motors some control authority in the z axis, see section II.A.1. The structure was also over built to limit resonant frequencies less than 10 Hz. Limiting any low vibration frequencies allowed the sensors to only measure the movement of the entire platform and not vibrations of the platform's structure which would induce error into the position and velocity measurements. Finally, the minimum acceleration of the CP/SADS about each axis should be  $0.1 \text{ rad/sec}^2$  with a

minimum velocity of 0.1 rad/sec upon startup. This requirement determined the inertia of the reaction wheels and was a factor in keeping the overall inertia of the platform down. The requirement was made to ensure substantial control authority over the platform.

### *1. Passive Balancing*

The structure was designed with balancing in mind so it is symmetric about the x- and y-axes. Refer to Figure 6 for a solid model of the CP/SADS. Hence, a requirement was imposed on the Structural Design team to insure the CP/SADSs center of mass could be adjusted in case components were shifted. The Structural Design team came up with two channels on each axis. Each channel contains a weight capable of moving along the channel. Therefore, the center of gravity of the CP/SADS can be adjusted. The center of gravity adjustment system is designed to make coarse adjustments to the center of gravity of the CP/SADS. The weights were made sufficiently large to apply a moment of 2 N-m.

### *2. Active Balancing*

While an active balancing system has not yet been fully developed, the structure of the CP/SADS was designed to accommodate the addition of such a system. An estimate of the mass of such a system was included in the solid modeling of the structural system and hence a location of the active balancing system was determined. The active balancing system of the CP/SADS is currently under development.

## **C. Actuator Design (Reaction Wheels)**

Four reaction wheels were designed to actuate the CP/SADS. For a complete analysis of the reaction wheel dynamics, see section II.A.1. This section outlines the sizing of the reaction wheels. In other words, the size and mass of the fly wheels was determined based on the imposed performance requirements.

Four Faulhaber 3863-24C DC Micromotors are used to actuate the platform. The motors are connected to the aluminum reaction wheels with inertias of 5.3552 Kg-cm<sup>2</sup>. The wheels trade off angular momentum by converting angular momentum of the wheels to angular momentum of the platform. The Faulhaber DC micromotors were used because of the high stall torque of 1.250 N-m. This is a substantial amount of torque for a 400 g motor and three times more than the equivalent servo motor. The stall torque is considerably more important for reaction wheel motors than the continuous torque, which is usually how motors are rated, because there is no external load applied to reaction wheel motors. The free spinning reaction wheels only produce a torque on the platform when the wheels are being accelerated and how fast the wheels can be accelerated is determined by the stall torque of the motors.

A large stall torque results in a large current draw from the motor. The stall torque is related to the applied current and the angular velocity of the rotor as,

$$T = K_M i - B\omega$$

where  $K_M$  is the motor torque constant, and  $B$  is the viscous friction of the motor. Here we assume the viscous friction of the motor is small and can be neglected.

The torque is the result of the large magnetic field produced by the current traveling around the motor windings. The motor windings naturally have a small amount of resistance. For

steady state conditions the applied voltage of a DC motor is equal to the counter electromagnetic force, otherwise known as back EMF, which results from the spinning of the motor added to the resistance times the current, or,

$$V = K_E \omega + Ri$$

The resistivity of the Faulhaber motor is  $0.62 \, \Omega$ . Since the maximum voltage is 24 V the start up current of the motor can be as high as 38.71 A. This is an enormous amount of current but this only occurs for a brief moment as the motor catches speed. It is expected that the motors will run on average not more than 10 A.

In order to achieve a minimal platform angular acceleration of  $0.1 \, \text{rad/sec}^2$  a large motor stall torque was needed. From Eq. (1) the CP/SADSs acceleration is determined by the motor stall torque and the overall platform's inertia. The motor stall torque can insure the platform's acceleration for a short time but the inertia of the reaction wheels are needed to determine how long this acceleration lasts and overall momentum transferred to the structure. By adding inertia to the reaction wheel it is possible to achieve the minimal platform angular velocity of  $0.1 \, \text{rad/sec}$ . The more inertia the wheels have the more momentum each wheel can achieve. Likewise an important factor in wheel momentum is the maximum velocity of the wheel, which in the case of the Faulhaber 3863-24C DC Micromotors is 6700 rpm. This is also known as the threshold speed because the motor reaches saturation at this speed and is no longer capable of supplying torque in one of its two directions.

It is important to note that the motor has two forms of braking. The motor can be braked by allowing the motor to coast so that the motor loses momentum by the friction in the bearings and the air friction due to the reaction wheel's spin. The second form of braking is called active braking or shunting and is done by allowing the two terminals of the motor to be left either both high or both low. This causes the motor to act as a generator and back EMF is sent through the Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) board and back to the motor. The MOSFET's resistance added to the motor winding resistance act to stop the motor's spin. Another important note is that the time it takes the motor to go from an initial speed to a final speed is approximately equivalent to the time it takes to go from the final speed and brake back to the initial speed. This means that the acceleration and deceleration of the motor have equivalent time constants.

Attached to the motor is an encoder. The IE2-64 encoder is a magnetic encoder with a resolution of 64 lines per revolution. The encoder gives frequency modulated pulses which are interpreted as the motor's speed.

#### **D. Power Subsystem**

There are four different power buses on the CP/SADS: 3 V, 5 V, 9 V, 24 V. All power comes from batteries onboard the CP/SADS and is distributed on the voltage regulation/data board (VRDB). There are 6 batteries in all: two 12 V and four 9 V. The 12 V batteries supply power to the motors while the four 9 V batteries provide power to the electronics.

The 12 V Power Sonic, 5.0 A-h, sealed lead acid batteries are connected in series to make the 24 V line to the MOSFET board. A lead acid battery was chosen over other batteries due to its large discharge rate at 10 times its capacity or 10 c, which is 50 A. This is much greater than Nickel Metal Hydride (Ni-MH) batteries or Lithium Ion (Li-I) batteries which have discharge rates of around 3 c and 0.5 c respectively. Though lead acid batteries tend to weigh more per

unit energy,  $Ah \cdot V$ , they still lead the way in discharge rate per unit mass due to their low internal resistance. Since total amount of energy in the battery is not as much of a concern as the amount of discharge current of the battery, the small lead acid battery is the obvious choice. The 50 A limit of the batteries is greater than the 40 A needed to run all four motors continuously.

The 24 V power bus is connected to four Pololu Motor Driver (MOSFET) Boards. Each Pololu driver acts as a voltage regulator for each motor. The Pololu driver has two n-fets and two p-fets which are pulsed to allow current to flow in both directions with minimal voltage loss.

The MOSFET board is controlled with two 5 V lines, IN 1 and IN 2 and one pulse width modulated 5 V line, EN (Enable). The control table, Table 1, shows the Boolean input for each line and its corresponding action.

Table 1: Motor Control Table

Enable	IN1	IN2	Action
H	L	L	Stop
H	L	H	Forward
H	H	L	Reverse
H	H	H	Stop
L	X	X	OFF

The VRDB is used as a central hub for all power and data lines with the exception of the 24 V motor power bus. There are three Ni-MH 9 V, 250 mAh batteries that power the VRDB.

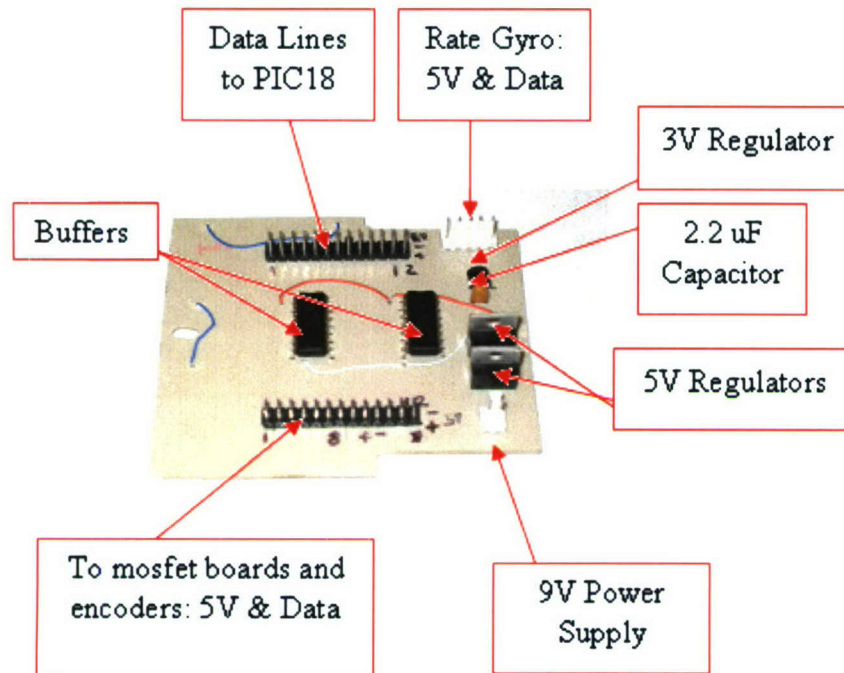
The 9 V batteries reduce to 5 V and 3 V buses through the use of voltage regulators on the VRDB. The 5 V bus powers the MOSFET boards (Pololu Drivers), rate gyros, buffers on the VRDB, and the encoders. The power allocation of each component on the CP/SADS is given in Table 2. A total of 377 mA is needed to run all components on the CP/SADS. All three batteries have a total energy of 750 mAh, which puts each battery well within the maximum drain rate of 1 c, 250 mA.

Table 2: Power consumption of each component

Component	Amperage (mA)	Number	Total Amperage (mA)
D200	80	4	320
Buffer	10	2	20
Rate Gyro	7	3	21
Encoder	4	4	16
Total			377

Figure 17 shows the VRDB and the various components and connectors. The VRDB was a custom made board etched using copper solvent solution. On board the VRDB is a pair of buffers that act to protect the Programmable Interrupt Controller (PIC), which is currently not incorporated into the PRWP, as well as increase the voltage of the output PWM and I/O lines from 3 V to 5 V. The Texas Instrument CD4050BE Buffers each have 6 Operational Amplifiers (Op-Amps). Together the pair of buffers has enough terminals for all 12 output ports. A 2.2  $\mu F$

capacitor is used to help steady the cyclic response of the 3 V regulator as well as absorb any transient spikes in the line. Capacitors were also added to the 5 V line for absorption purposes.



**Figure 17: Voltage Regulation/Data Board (VRDB)**

### E. Sensors

Attitude determination of the CP/SADS is accomplished by three position and velocity sensors on board. The rate gyros and IMU are both dead reckoning sensors which determine position and velocity by forces that act on them. The third sensor, the star tracking camera, determines position and velocity by means of a reference point such as a star in a fixed reference frame. Over time any dead reckoning sensor will accumulate errors. As these errors grow the actual position and the sensed positions diverge. This phenomenon is known as drift. A star tracker is often used to update the known attitude of the spacecraft periodically. Attitude maintenance or upkeep is usually done on a half-day basis depending on the satellite.

Three Silicon Sensing CRS03-01 Rate Gyro sensors are placed on each of the three body axes to record angular rate measurements of the rotating platform. The rate gyros used are Micro-Electrical to Mechanical System (MEMS) devices. The rate gyros work using a Silicon Vibrating Structure Gyroscope (Si-VSG) which measures vibrations of an etched silicon ring in a magnetic field. Current is passed through the conductors on the ring to keep the ring vibrating at a constant 14 kHz.<sup>6</sup> This closed loop system produces a Rateout Voltage ( $V_O$ ) signal that is proportional to the angular rate being sensed. Determining angular position about each axis can be accomplished by integrating the voltage signal with respect to time. This process for determining angular position is less accurate over time as errors in the signal accumulate.

When integrated, the IMU can be used in conjunction with the rate gyros to give a more accurate position as time increases. The LN-200 has a drift rate of one degree per hour. The low drift rate of the LN-200 is due to the ring-laser gyroscope which measures the interference patterns of light emitted in opposite directions. The main cause of drift is due to thermal expansion problems which change the length the paths of light travel. The advantage over

MEMS devices is the lack of moving parts which means no need for calibration other than the need for a zero-point adjustment.<sup>7</sup>

For a star tracker the ZTV Micro-Camera was chosen mostly for its small size, light weight, and refresh rate of 30 Hz which is fast enough for its application. The purpose of the camera is to locate the position of an LED placed above the platform. By tracking the LED's motion the platform's relative position can be determined relative to the inertial frame of the LED. With the camera attitude maintenance of the platform can be scheduled periodically to alleviate the error buildup of the rate sensors (the ZTV camera is not currently integrated)

#### **F. Data Acquisition**

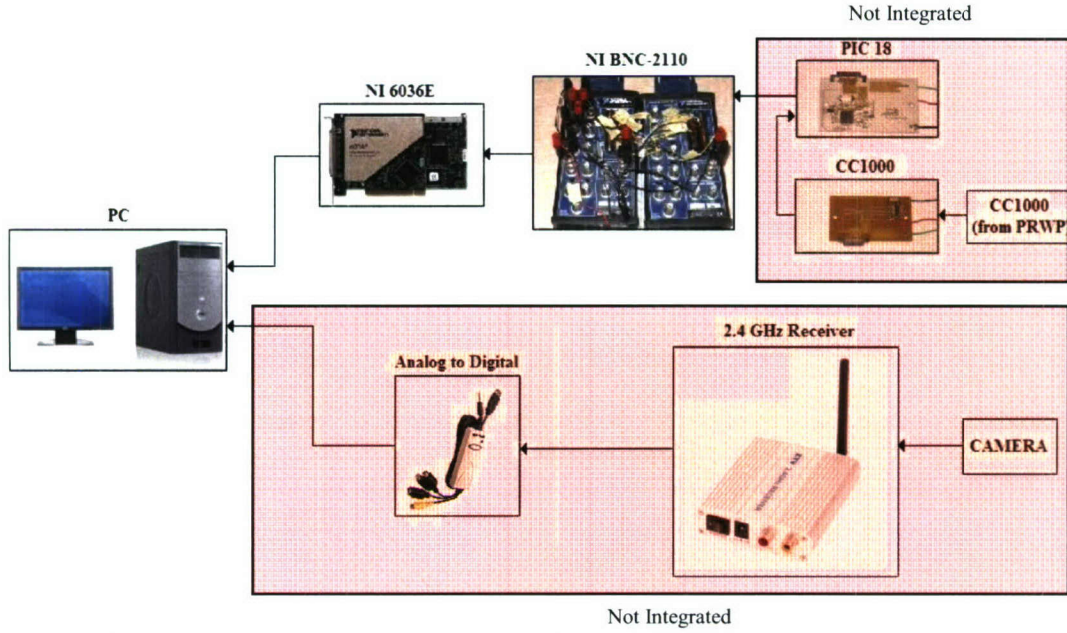
One of the main objectives for future versions of the CP/SADS is to remove the need for an umbilical connection. In order to achieve this goal all data needs to be received and transmitted wirelessly from the platform to the ground station and vice versa. Two transmitters will be used to send data from the platform to the ground. The first transmitter, which has already been installed on the platform, is embedded in the camera and it sends data to the ZTV 2.4 GHz Receiver positioned on the ground.

The ground based 2.4 GHz Receiver supplies an analog RCA output. The RCA output is converted into a digital signal by the Kworld analog-to-digital adapter that can be connected via a USB port to the PC. Using Matlab's image acquisition and processing toolboxes it is possible to process the data and find the position of the platform with respect to the inertial frame.

The second transmitter, which will be installed in a future version, is the Chipcon CC1000 which acts as a receiver and transmitter, otherwise known as transceiver. The CC1000 lies on a Prototype Circuit Board (PCB) that also holds the PIC 18. Data is captured by the PIC's A/D, digital I/O, and counter ports. The data is then sent to the transceiver where it is stored. The stored data is sent in packets to another ground based CC1000 transceiver. Likewise data can be sent from the ground to the platform if the process is reversed (For more information on the design and software of the CP2 Transceiver see Huerta<sup>8</sup>).

Currently the umbilical connection from the platform ties into a pair of National Instruments BNC-2110 connector boxes. The boxes individually hold two 16-bit analog outputs, sixteen 16-bit analog inputs, eight digital I/O lines and two 24-bit counters. The BNC-2110s connect to two PCI 6036E cards installed in the computer. The PCI cards are also made by National Instruments and work to interface with Matlab using Matlab's Real-Time Windows Target.

Figure 18 shows the proposed ground based hardware schematic. Data is received through both the ZTV 2.4 GHz receiver and the CC1000 transceiver and sent to the PC (The current version of the CP/SADS does not incorporate the CC1000 transceiver due to the many intricacies of the communication protocol; instead an umbilical cord is attached that connects the VRDB to the National Instruments BNC-2110 breakout board). Matlab works both to process the data and to close the loop by commanding the appropriate shunt and drive voltages as well as the direction polarity for the motors onboard the CP/SADS.



**Figure 18: Ground based hardware schematic**

### G. Full State Feedback Controller Design

Since the kinematics and dynamics of the platform are nonlinear it was important to first linearize the system to derive a control law using full-state feedback. Asymptotic stability of the nonlinear system is shown later using a Lyapunov function. The equations of motion of the platform are given by Eqs. (20) and (29). Linearizing the vectors  $\vec{q}$ ,  $q_4$ , and  $\vec{\omega}$  about  $q_4 = 1$  gives,

$$\vec{\omega}_p \approx \delta \vec{\omega}_p$$

$$\vec{q} \approx \delta \vec{q}$$

$$q_4 \approx 1 + \delta q_4$$

which results in the linearized equations of motion,

$$\vec{M} \approx I_M \delta \ddot{\vec{\omega}}_p = T_C + \vec{M}_{Ext}$$

$$\delta \dot{\vec{q}} \approx \frac{1}{2} (\vec{\omega}_p) \quad (30)$$

$$\delta \dot{q}_4 \approx 0$$

Solving for  $\vec{\omega}_p$  and  $\dot{\vec{\omega}}_p$  results in,

$$\delta \vec{\omega}_p = 2 \delta \dot{\vec{q}}$$

$$\delta \ddot{\vec{\omega}}_p = 2 \delta \ddot{\vec{q}}$$

Applying the control torque,

$$\vec{T}_c = -K\vec{q}_e - C\vec{\omega}, (31)$$

and assuming no external moment is applied to the platform results in,

$$I_m \delta \dot{\vec{\omega}} + C \delta \vec{\omega} + K \delta \vec{q}_e = 0.$$

Substituting Eq. (3) into the dynamic equation gives,

$$\frac{d^2 \delta \vec{q}_e}{dt^2} + CI_M^{-1} \frac{d \delta \vec{q}_e}{dt} + \frac{1}{2} KI_M^{-1} \delta \vec{q}_e = 0. (32)$$

Relating Eq. (32) to the damping,  $\zeta$ , and natural frequency of the system,  $\omega_n$ , gives the control matrix constants K and C,

$$K = 2I_m \omega_n^2 = \begin{bmatrix} 2I_1 \omega_n^2 & 0 & 0 \\ 0 & 2I_2 \omega_n^2 & 0 \\ 0 & 0 & 2I_3 \omega_n^2 \end{bmatrix} \quad (33)$$

$$C = 2I_m \zeta \omega_n = \begin{bmatrix} 2I_1 \zeta \omega_n & 0 & 0 \\ 0 & 2I_2 \zeta \omega_n & 0 \\ 0 & 0 & 2I_3 \zeta \omega_n \end{bmatrix}$$

The inertia matrix,  $I_m$ , of the platform is approximately a diagonal matrix due to the symmetry of the platform. For this reason only the moments of inertia were taken and all products of inertia were omitted. However, movement of the counter weights on the platform can change this symmetry. For cases when the platform is not especially symmetric the principal moments of inertia can be found and a transformation matrix can be used to transform the control torques to the body axes.

A full picture of the simulation can be seen in Figure 18. Each subsystem is represented by equations derived earlier in this report. A gravity torque can be introduced but was not since the simulation was only taken about the z body axis due to the difficulty of making the platform neutrally stable with the large counter weights. A normalizing block was put into the quaternion loop because as the simulation progresses the quaternions tend to drift.



From Wie<sup>1</sup>, asymptotic stability of the platform over a wider range of motion can be shown with the control law from Eq. (31) with gains defined from Eq. (33). First, generate a Lyapunov function of the form,

As shown in Wie<sup>1</sup>, a asymptotic stability is achieved for gains chosen such that,

where  $k_1, k_2$ , and  $k_3$  are the diagonal terms in the K matrix from Eq. (33). The gains for the C matrix from Eq. (33) must simply be positive constants.

$$\begin{aligned} V(x) &\geq 0 \\ \dot{V}(x) &\leq 0 \\ \ddot{V}(x) &\text{ is bounded} \end{aligned}$$
$$\dot{V}(x) = -\vec{\omega}_p^T C K^{-1} \vec{\omega}_p \geq 0$$

The second criterion is satisfied because the control gains  $C$ , and the inverse of  $K$  are positive definite matrices. The angular velocities of the platform are squared so this term is also positive. The second derivative of the Lyapunov function is given by

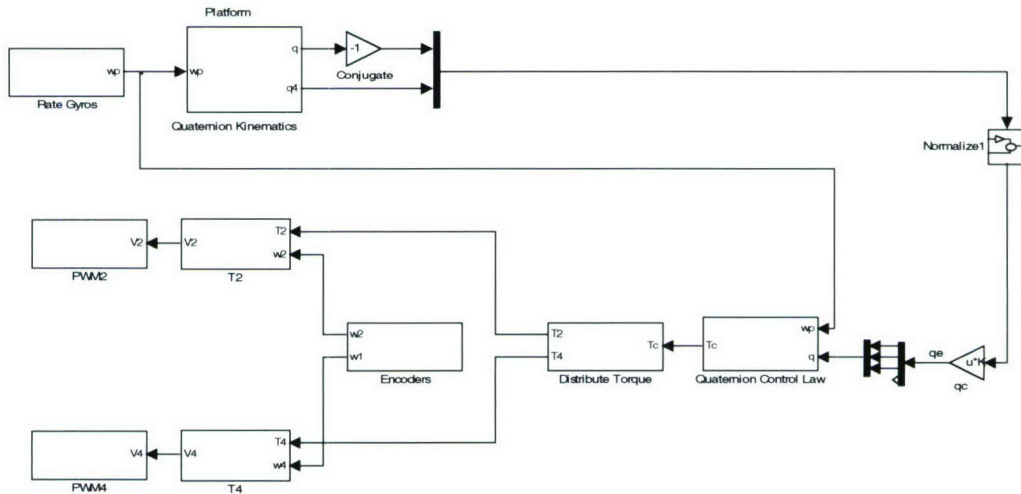
$$\ddot{V}(x) = -2\dot{\omega}_p^T C K^{-1} \dot{\omega}_p.$$

Barbalat's Lemma (criterion #3) is used to show asymptotic stability. The second derivative of the Lyapunov function is bounded meaning the Lyapunov derivative diminishes to zero as time approaches infinity, i.e.,  $\dot{V}(x) = 0$  as  $t \rightarrow \infty$ . This means the angular velocity of the platform

diminishes to zero as time approaches infinity,  $\dot{\omega}_p = 0$  as  $t \rightarrow \infty$ . Therefore, the equations of motion of the platform in conjunction with the control law from Eq. (31) are asymptotically stable.

### I. MatLabReal Time Control Toolbox

To test the CP/SADS a control command was designed to move the platform through an angle of  $-90^\circ$  about the Z-axis. For this maneuver only two reaction wheels were used because the encoder on one of the motors was damaged during construction of the reaction wheel flywheels. For this reason the real time controller shown in Figure 20 differs from the simulated controller shown in Figure 19 in that there are only two motors (number 2 and number 4) instead of 4.



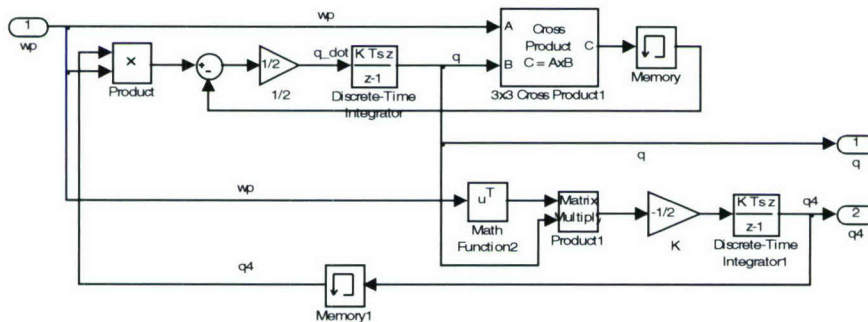
**Figure 20: Real Time PRWP Controller**

The real time controller is different from the simulated system since the real time system is based on digital data acquisition and command. The motor controllers use pulse width modulation to control motor speed. The rate gyros from Silicon Sensing were used to acquire body rates. The manufacturer specified that the angular rate ( $R_a$ ) is given by,

$$Ra = \frac{-\frac{1}{2} \times V_{dd} + V_o}{\left( SF \times \frac{V_{dd}}{5} \right)} \text{deg/sec.} \quad (34)$$

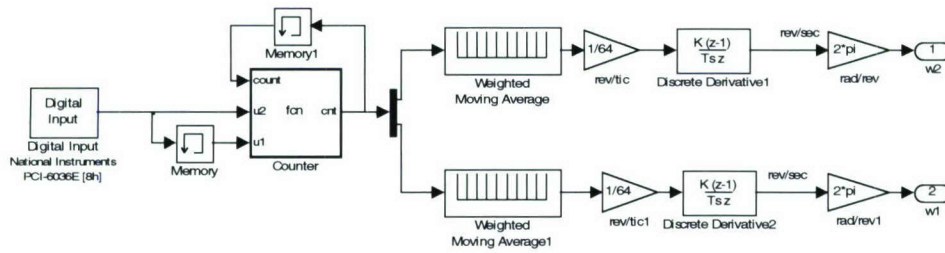
The voltages  $V_0$  and  $V_{dd}$  represent the output voltage of the rate gyro and the reference voltage supplied to the rate gyro respectively. The scale factor (SF) relates the difference in voltage to the angular rate in deg/sec. Figure 21 shows the rate gyro subsystem. The analog inputs represent the z axis gyro and the reference voltage. From Equ. (34), the voltages are converted into a z axis rate in deg/sec. This rate is then converted into rad/sec. A bias is added to the z axis rate to give a more accurate measurement. The bias is modified over time as the rate gyro heats up, but not during a command to the CP/SADS To determine the correct bias the platform is put on a table and held stationary. A ten second sample is taken and integrated. A bias is then inserted to try to minimize the drift rate to zero. Without a bias the platform may drift as much as 0.3 deg/sec. Upon insertion of a bias the drift rate can be as low as 0.05 deg/sec (It is good to wait about twenty minutes to allow the gyro to come to thermal equilibrium before inserting a bias). The X and Y axis rates are set to zero and incorporated as one body rate. The rates are reversed because of the sensor orientation.

The quaternion kinematic subsystem shown in Figure 21 has only two changes from the quaternion subsystem from the simulated system. Two memory blocks are inserted into the feedback because Matlab is unable to perform an algebraic loop without a delay. In the simulated controller the model was running in continuous time but with the real time controller all models are running in discrete time.



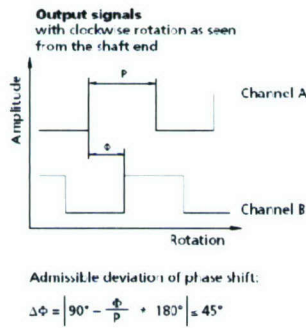
### Figure 21: Quaternion Kinematics Subsystem

Figure 22 shows the encoder subsystem which converts each motor's encoder signals into the corresponding velocity of each motor in rad/sec. The counter embedded function looks for the rising edge of a pulse and then increments the count or decrements the count if the motor is reversed.



**Figure 22: Encoders Subsystem**

Since there are two encoders per motor (see Figure 23) direction is determined by relating channel A to channel B. Positive rotation according to Figure 11 is the opposite of that shown in Figure 23 so the output is reversed in the embedded code.



**Figure 23: Encoder output**

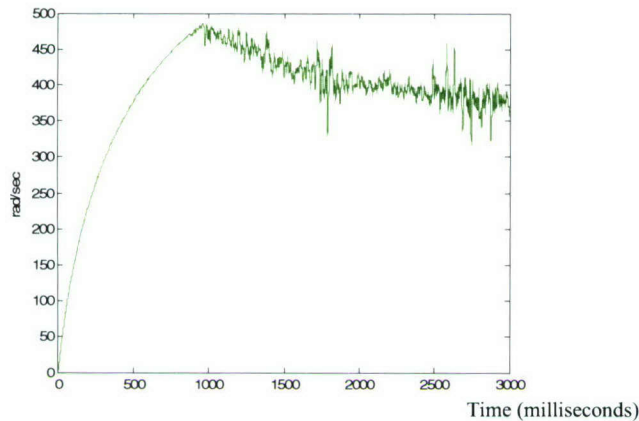
In the encoder subsystem the counts for each motor are separated and then a running average is taken to both smooth out the signal and delay the signal from 20 kHz to 100Hz. The counts are then divided by 64 because there are 64 counts in one revolution. The smoothed signal is then feed through a discrete time derivative block which gives the wheel's velocity in rev/sec. A gain then converts the wheel velocity to rad/sec.

The real time controller is limited to sampling at a rate of 20 kHz before memory buffer overflow problems occur. Due to this limit there exists a threshold on the detection of the wheel velocities. Past a wheel velocity of 490 rad/sec aliasing of the wheel occurs. This condition is called the Nyquist criterion, and is equivalent to saying that the sampling frequency must be strictly greater than twice the signal's bandwidth,

$$\frac{20000(\text{samples})}{\text{sec}} \cdot \frac{\text{rev}}{64(\text{period})} \cdot \frac{\text{period}}{4(\text{samples})} \cdot \frac{2\pi(\text{rad})}{\text{rev}} = 490.8750 \left( \frac{\text{rad}}{\text{sec}} \right)$$

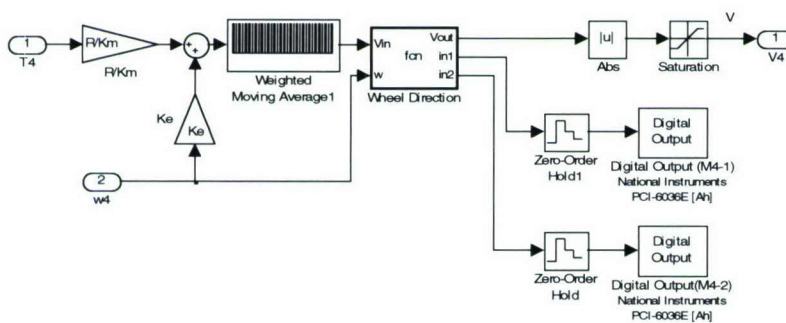
There are 4 samples in each period due to the two encoder channels A and B. If direction was not being accounted for then the wheel velocity detection threshold would be 982 rad/sec; this is a faster speed than the motor can achieve. Since the direction of the motors is important the

motor speeds were limited to under 490 rad/sec. Figure 24 shows a motor accelerating to 490 rad/sec, after which aliasing takes place and the wheel speed becomes sporadic.



**Figure 24: Aliasing of the reaction wheel**

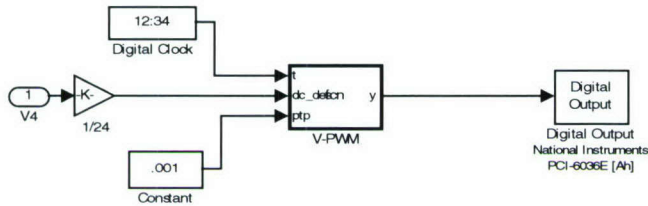
Figure 25 shows the direction and torque subsystem. This subsystem converts the torque and the velocity of the wheel for each motor into the appropriate spin direction and voltage. The voltage to the motor is found by relating the commanded torque to an equivalent commanded voltage. The commanded voltage is found by multiplying the torque by the motor resistance ( $R$ ) divided by the motor torque constant ( $K_m$ ) and adding it to the voltage created by the spinning armature given by multiplying the motor spin by the back EMF constant ( $K_e$ ). A running average of the voltage is taken to smooth out the voltage signal. The wheel direction embedded function takes the applied voltage and motor velocity as input and gives the motor direction as output by providing a high or low input to the digital outputs that head to pins IN1 and IN2 of the MOSFET board. The wheel direction embedded code protects the motor from reversing polarity of the motor terminals when the motor is spinning fast, which would cause a current spike that could destroy the motor coils. The motor is protected at high speeds by first shunting the motor before reversed in direction commanded to the motor. The voltage,  $V_{out}$ , from the wheel direction embedded function is then fed through an absolute value block and a saturation block to make sure the voltage,  $V_4$ , to the motor is both positive and no greater than the max voltage of 24 V.



**Figure 25: Direction and Voltage Torque Subsystem (T2-T4)**

Figure 26 shows the voltage to pulse width modulation (V-PWM) subsystem. The V-PWM subsystem takes the voltage from the direction and voltage torque subsystem ( $V_4$ ) and converts the voltage into an equivalent PWM signal. Inside the V-PWM

embedded function, which takes the clock running at 20 kHz, a voltage which is limited between 0 (no voltage) and 1 (24 V), and the pulse width period (ptp) at .001, as inputs. The V-PWM gives a pulse width modulated signal at 1 kHz as output to the motors. The PWM signal is fed to the enable pin on the MOSFET board which controls the motor's voltage from 0 to 24 V. The duty cycle of the PWM channel can only be incremented by 5% because 5 is the least common integer value between the 1 kHz PWM frequency and 20 kHz output frequency of the digital output.



**Figure 26: Voltage to Pulse Width Modulation Subsystem (PWM2-PWM4)**

## IV. Experimental Results of the Cal Poly Spacecraft Attitude Dynamics Simulator

### A. Simulator Test Setup

The CP/SADS was initially tested with two different maneuvers. First a 90 degree rotation about the z-axis of the CP/SADS was performed. The goal was to show that the sensors and actuator were performing adequately and to demonstrate the initial point capability of the CP/SADS. The second test involved disturbing the CP/SADS with an external torque and commanding the CP/SADS to return to its original orientation. This test demonstrates the ability of the CP/SADS to respond to external torques and allows for the study of the motor saturation characteristics. The CP/SADS was set up with the real time controller as described in the above sections for both tests.

### B. Test 1: Rotation about a Single Axis

Two tests were conducted on the PRWP using the real time and simulated controllers to show controllability of the platform. The first test required the CP/SADS to maneuver  $-90^\circ$  about the z axis. The initial quaternion for the CP/SADS was set to

$$\vec{q} = \begin{bmatrix} 0 & 0 & \sqrt{2}/2 \end{bmatrix}^T$$

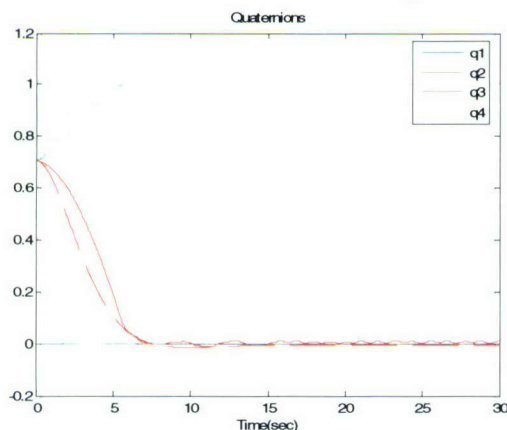
$$q_4 = \sqrt{2}/2$$

which is a  $90^\circ$  rotation around the Z axis in the inertial frame. The commanded quaternion was set to  $\vec{q}_C = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$  and  $q_{C4} = 1$  which is  $0^\circ$  in the inertial frame. The rotation between the initial and commanded quaternions is a  $-90^\circ$  maneuver in the z body axis.

For the simulated controller, K and C gains from Eq. (31) were found that command the platform to its final position quickly and with little overshoot. These same K and C gains were used in the real time controller but altered slightly to match the simulated controller. Differences

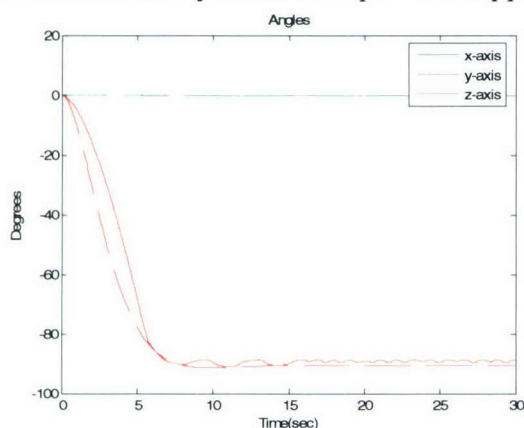
in the gains of the two controllers are due to two factors. The first is that there are differences between the actual and simulated inertias and the second is that the real time controller is not fully modulated like the simulated controller. The differences between full modulation and finite pulse width modulation are especially pronounced in the small angle steady state maneuvers after the platform has maneuvered itself to its commanded position.

Figure 27 shows the real quaternions of the platform compared to the simulated quaternions (dashed lines) for the first test where the platform maneuvers itself to  $-90^\circ$ . The K gain for the simulated controller was set to 1.1 whereas the real time controller had a K gain of 0.9 making the real time controller slightly more sluggish at the beginning of the maneuver. The C gain for the simulated controller was set to 0.2 whereas the real time controller had a C gain of 0.3 causing the real time controller to dampen out faster. It is also speculated that the current draw from the batteries was saturated during the initial part of the maneuver.



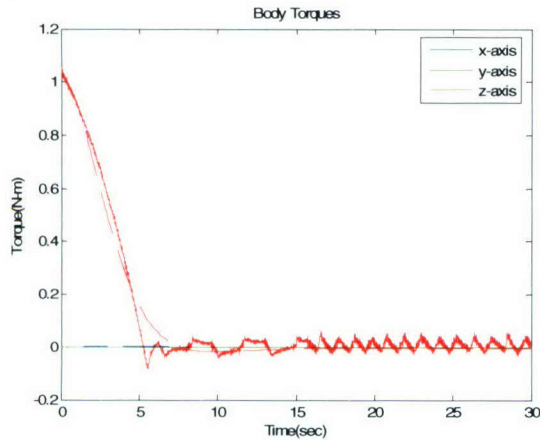
**Figure 27: Quaternions of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 28 shows the corresponding body angles, which are the same as the Euler angles. Since the Z inertial axis is in line with the z body axis the angles in Figure 28 correspond to both the inertial and body axes. The platform approaches  $-90^\circ$  in about 7 seconds.



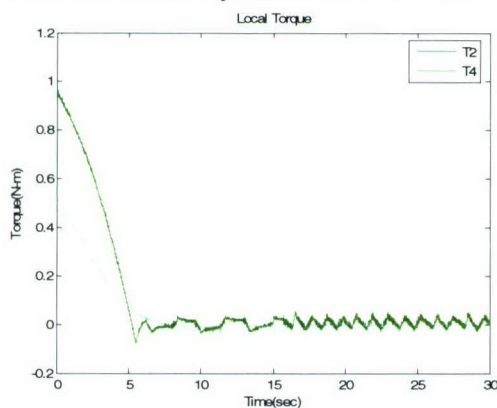
**Figure 28: Euler angles of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 29 shows the body commanded torques from the quaternion control law subsystem. The x and y axes torques are zero because the platform is only maneuvering in the z body axis. Since the real time controller is not fully modulated the real time torques after the platform has reached its commanded position are shifted back and forth meaning the real wheel speed needed to maintain this position is somewhere between the incremented wheel speeds. This phenomenon is known as clocking. The platform compensates for this by switching torques in rapid procession.



**Figure 29: Body torques of the simulated platform (dashed lines) and actual platform (solid lines)**

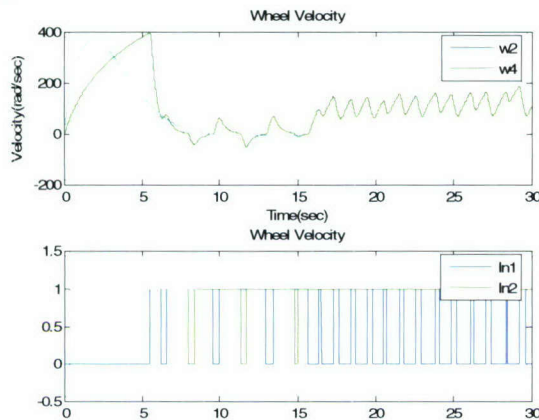
Figure 30 shows the local wheel torques applied to each wheel from the distributive torque subsystem. Since rotations were only in the z body axis all four local torques of the simulated controller are equal as well as both local torques of the real time controller. The real local torques are twice the simulated local torques. The multiplier of two adjusts the real time controller since it only has half the number of reaction wheels as the simulated controller.



**Figure 30: Local wheel torques of the simulated platform (dashed lines) and actual platform (solid lines)**

Wheel velocities for the simulated and real platform can be seen in Figure 31. The wheel velocity for the simulated platform starts at zero and after the maneuver is accomplished the ending wheel velocity is again zero. The simulated controller assumes no friction therefore momentum is exchanged directly to the platform. Since there is neither loss in energy nor any external torques the momentum from the platform is transferred back to the wheel to eventually

stop the platform. This momentum is equal to the momentum it took to accelerate the platform which causes the ending velocity of the wheel to be zero. Since the real platform does have friction, both in the air bearing and in the reaction wheel (due to the air), the final velocities do not go to zero but are left at around 80 rad/sec.

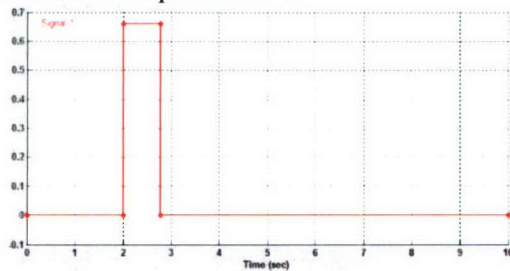


**Figure 31: Wheel velocities (top) of the simulated platform (dashed lines) and actual platform (solid lines) and direction inputs to the MOSFET board for the actual platform (bottom)**

### C. Test2: Disturbance Response

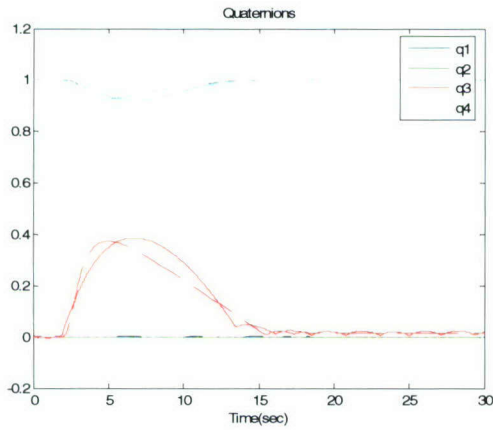
A second test was conducted to show platform stationary maintenance ability with the introduction of a disturbance. Both initial and commanded quaternions were set to

$\vec{q} = [0 \ 0 \ 0]^T$  and  $q_4 = 1$  for this test. In the second test an external moment was applied to the platform to see how the platform would respond. This external moment for the real time controller came in the way of a slight hand nudge in the z body frame whereas the external moment for the simulated controller tried to mimic this hand nudge using the square waveform shown in Figure 32. The real time controller test was completed first and the simulated controller test was designed to imitate the real time controller test by changing the amplitude and period of the square wave.



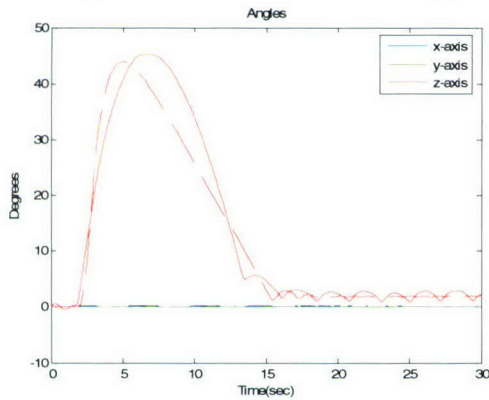
**Figure 32: External moment applied to the simulated platform in the z body axis intended to match moment applied to the actual platform.**

Figure 33 shows the quaternions of the real and simulated controllers when an external moment is applied in the z body axis to the platform 2 seconds into the test.



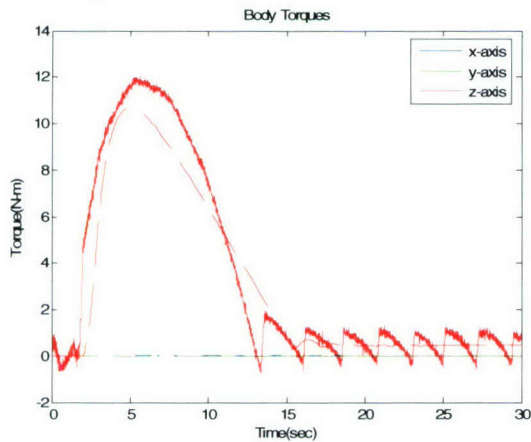
**Figure 33: Quaternions of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 34 shows the corresponding Euler angles of the real and simulated platforms. The platform is pushed about  $45^\circ$  off its original position before it recovers.

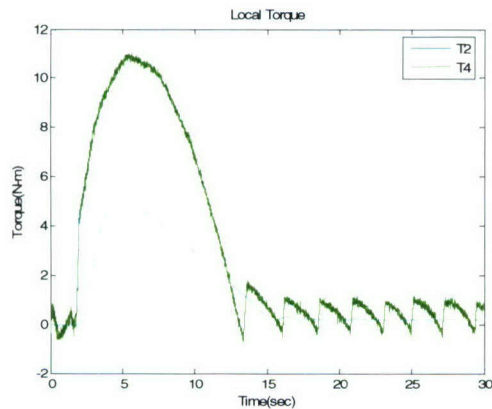


**Figure 34: Euler angles of the simulated platform (dashed lines) and actual platform (solid lines)**

Figures 35 and 36 show the body and local commanded torques of the platform. The real time controller local torques are twice that of the simulated torques because the real time platform only modulates two wheels instead of four on the simulated platform.

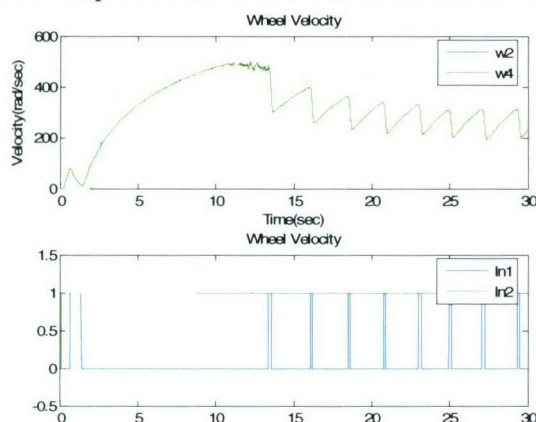


**Figure 35: Body torques of the simulated platform (dashed lines) and actual platform (solid lines)**



**Figure 36: Local wheel torques of the simulated platform (dashed lines) and actual platform (solid lines)**

Wheel velocities for the real and simulated platform are shown in Figure 37. Unlike the first test, the second test of the platform applies an external moment. The wheels increase velocity to combat the positive moment applied. The opposing moment the wheels produce causes the platform to decelerate and eventually accelerate in the opposite direction to get back to its original orientation. Upon arriving at its original position some of the momentum from the platform is transferred back to the wheels but due to the external moment which was applied to the platform the wheels need to keep a constant final velocity. The energy in the spinning wheels at the constant final velocity is proportional to the energy it took to nudge the platform from its initial position. If the platform were given a greater external moment or a series of successive external moments the reaction wheels would reach a limit in which they could no longer produce the opposing torque needed to combat the external torque. When this happens the reaction wheel is said to have saturated. The reaction wheel's saturation speed is the motor's maximum speed which is 710 rad/sec for the Faulhaber motors.



**Figure 37: Wheel velocities (top) of the simulated platform (dashed lines) and actual platform (solid lines) and direction inputs to the MOSFET board for the actual platform (bottom)**

## V. Conclusions, Recommendations and Future Work

The initial development of the CP/SADS has been successful. Both tests described above demonstrate the ability of the CP/SADS to perform as a spacecraft simulator about a single axis. The first test demonstrates the pointing accuracy of the CP/SADS. The pointing accuracy of the CP/SADS needs to be fully characterized, but appears to be around 2-3 degrees.

The second test verifies the well known phenomenon which occurs with momentum exchange devices. Namely, external moments can cause reaction wheel saturation. While the applied external moment was not precise, the concept is surely demonstrated.

The initial development of the CP/SADS has proven to be a very useful learning experience for both students and faculty involved with the project. From designing subsystem hardware to writing real time control algorithms, it was the first time many of these tasks were performed by any of the team members involved. The development of the CP/SADS has provided an invaluable set of institutional knowledge for the students and faculty at Cal Poly.

There are several obvious required improvements before the CP/SADS is truly ready for space system simulation and testing. To complete the development of the CP/SADS, the following items are considered high priority:

- Design and manufacture of the Fine Balancing System
- Integration of the wireless data acquisition and commanding system
- Integration of the Northrop Grumman LN-200 Inertial Measurement Unit

Before the CP/SADS can emulate a fully agile spacecraft, the CP/SADS must be balanced about all three axes. The Fine Balancing System, encompassing the Center of Gravity System Identification algorithm and the Fine Balance Mass Tuning System is critical to collocating the CP/SADS center of gravity and the pivot point of the spherical air bearing. A considerable effort involving many students and faculty is currently underway.

To further enhance ability of the ability of the CP/SADS to emulate the microgravity environment of space, the umbilical cord of wires used to command and control the CP/SADS has to be removed. Hence, the CP/SADS has to move to a wireless data acquisition system. Several hardware components have been identified to design and build such a system. Several students and faculty from the Computer Engineering Program within the Computer Science Department at Cal Poly have been brought on board to work with students from the Aerospace Engineering Department to complete the wireless data subsystem project.

With a balanced simulator which has a wireless data acquisition system, the CP/SADS will be ready for the integration of the Northrop Grumman LN-200 Inertial Measurement Unit. The LN-200 is critical to the CP/SADS since it is capable of independently providing both angular and angular rate measurements. The LN-200 is a much more precise instrument than the current set of MEMS rate gyros used by the CP/SADS. The LN-200 also contains a set of accelerometers which can be used to find angular accelerations of the CP/SADS without taking derivatives of the angular velocity. The angular acceleration is used to identify the center of gravity and inertial characteristics of the CP/SADS.

Once the three items listed above are complete, the CP/SADS will be ready for more projects and tests. The CP/SADS will be capable of simulating various control laws, testing other sensors and actuators, and testing conceptual designs for future satellite missions. The goal of the CP/SADS platform is to provide a test bed for future satellite attitude control algorithms, sensors and actuators. Our group intends to use the CP/SADS to demonstrate fine pointing control of

micro-class satellites for use with optics based systems. We are currently in the conceptual design phase of a small space based optical telescope.

Our group also intends to focus on several areas of fruitful areas of research and engineering design and development including:

- Adaptive Disturbance Rejection of Tonal Reaction Wheel Assemble Disturbances
- IMU, Rate Gyroscope and GPS Sensor Fusion for Attitude Determination and Control
- Linear and Nonlinear System Identification of Reaction Wheel Systems
- Line-of-sight Geolocation for Remote Sensing Space Systems
- Autonomous Control of Satellite Formations
- PolySat Rate Gyro Characterization and Calibration
- Coupled Reaction Wheel and Control Moment Gyroscope Attitude Control

Each of these topics is an exciting area of current research or interest to both the academic and industrial community. By providing an environment for students to learn cutting edge satellite navigation and attitude control techniques, Cal Poly will help our students become technical leaders in industry upon graduation. The CP/SADS and its associated laboratory will also provide academic and industry partners a place to test, develop and integrate future space based navigation and control components.

## **VI. Acknowledgements**

This work was sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855.

## **VII. Bibliography**

1. Wie, Bong, *Space Vehicle Dynamics and Control*, Reston, VA, 1998.
2. Christian, John, Turbe, Michael, Kabo, Eric, Manno, Laura, and Johnson, Eric, "Development of a Variable Inertia Reaction Wheel System for Spacecraft Attitude Control," *AIAA Guidance, Navigation and Control Conference*, AIAA-2004-5132, Providence, Rhode Island, August 2004.
3. Rowell, Derek, Wormley, David N., "State Equation Formulation," *System Dynamics*, New Jersey, Prentice Hall, 1997.
4. Healy, Patrick, "Mass Property System Identification of a Spacecraft Simulator," California Polytechnic State University, San Luis Obispo, Master's of Science Thesis, 2006.
5. Martin, Joe, Quinto, Adrian, Nash, Amanda, "Reaction Wheel Platform (Satellite Simulator Structure)," California Polytechnic State University, San Luis Obispo, Senior Project, 2006.
6. Pearce, C, "The Performance and Future Development of a MEMS SIVSG and its Application to the SIIMU," AIAA 2001-4410, 2001.
7. Wiesel, William E., "Rigid-Body Dynamics," *Spaceflight Dynamics*, 2<sup>nd</sup> ed., Irwin McGraw-Hill, Massachusetts, 1995.
8. Huerta, Christopher, Derek, "Development of a Highly Integrated Communication System for use in Low Power Space Applications," California Polytechnic State University, San Luis Obispo, Master's of Science Thesis, 2006.
9. Mehiel, Eric A., and Balas, Mark J. "Nonlinear Direct Model Adaptive Control", *Submitted, Journal of Guidance, Dynamics, and Control*.

10. Space Systems Simulation Laboratory, Department of Aerospace Engineering, Virginia Polytechnic Institute and State University, January 2006.  
<<http://www.sssl.aoe.vt.edu/dsacss/index.php>>.
11. Tanygin, Sergei, and Trevor Williams. "Mass Property Estimation Using Coasting Maneuver." *Journal of Guidance, Control, and Dynamics*, Jul. - Aug. 1997: 625-632.

#### **VIII. Publications and Presentations Related to Project**

##### **Full Publications**

Copies of full publications follow.

# The Cal Poly Spacecraft Attitude Dynamics Simulator

Eric A. Mehiel\* and Carson O. Mittelsteadt  
*California Polytechnic State University, San Luis Obispo, CA, 93402*

The Cal Poly Spacecraft Attitude Dynamics Simulator (CP/SADS) is a four reaction wheel, pyramid configuration momentum exchange device which simulates the attitude dynamics and kinematics of a rigid body in space. The structure of the reaction wheel pyramid rests on a spherical air bearing. When the pivot point of the bearing is aligned with the center of gravity of the pyramid structure, no gravity torque is incumbent on the simulator. The combination of no gravitational torque on the structure and the four wheel momentum exchange device allows for dynamically realistic simulations of the attitude dynamics and kinematics of spacecraft in earth orbit. The goal of the CP/SADS project is to develop a test bed for spacecraft attitude control development, actuator and sensor design and characterization, inertial navigation sensor and algorithm development, and precision pointing algorithm development. This paper outlines the design components of the CP/SADS which are unique. Specifically, the CP/SADS uses a four axis pyramid configuration instead of the more typical orthogonal three axis configuration. The CP/SADS employs onboard power and will soon take advantage of wireless data transmission.

From the beginning of the project, performance requirements were internally developed and imposed. The CP/SADS project is also inherently multidisciplinary and students with expertise from departments outside the Cal Poly Aerospace Engineering department were utilized. However, the CP/SADS project was managed by students and faculty from the Aerospace Engineering department. The structure of the simulator was designed and built by Cal Poly Mechanical Engineering Students. The reaction wheels, all supporting electronics, and software required for operation, were designed and developed by Cal Poly Aerospace Engineering graduate students and Cal Poly faculty.

The simulator preliminary performance results are presented and discussed. Currently the simulator provides motion in only one axis since the final balancing of the system has not been performed. The technique used to balance the simulator in all three axes is still in development. Two tests were conducted on the simulator to show its performance capabilities. The first test demonstrated the capability of the simulator to maneuver from arbitrary initial position to some desired final position. The second test showed the ability of the simulator to maintain a fixed position when perturbed. The results of the platform correlated well with the predicted results found through analytical analysis and simulation. The test results were also used to verify the performance requirements of the CP/SADS were met, including a required maximum angular velocity of 0.1 rad/sec and a maximum acceleration of 0.1 rad/sec/sec.

---

\* Assistant Professor, Aerospace Engineering Department, 1 Grand Ave, AIAA Member

## **I. Spacecraft Attitude Dynamic Simulators and Momentum Exchange Devices**

### **A. Introduction**

THE primary purpose of a satellite attitude control system is to stabilize and to provide attitude control for the spacecraft. Momentum exchange devices are one method commonly used to provide the control torques necessary for accurate attitude control. One type of momentum exchange device commonly employed is known as a Reaction Wheel. Reaction Wheels operate by exchanging an equal and opposite amount of angular momentum between the fly wheel of the Reaction Wheel and the system. Reaction Wheels exchange angular momentum with the body of the spacecraft by changing the angular velocity of the fly wheel. Hence, the direction of the angular velocity vector remains constant while the magnitude of the vector changes. When several Reaction Wheels are combined onto one spacecraft platform, the changes in the different wheel speeds are combined to affect the attitude of the spacecraft in all three axes. In general a fully agile spacecraft needs three reaction wheels at a minimum and sometime four are used to provide redundancy of the attitude control system.

Reaction Wheels differ from other Momentum Exchange devices such as Control Moment Gyroscopes, CMGs, in design but not in principle. Control Moment Gyros exchange momentum with the spacecraft by changing the direction of the angular velocity vector of the flywheel while keeping the direction constant. The dynamic analysis of Reaction Wheels is typically easier since the Inertia Matrix of the system does not vary with time since Reaction Wheels are typically designed to be axisymmetric. In contrast, CMGs momentum exchange devices usually cause the inertia matrix of the system to be time varying. There are also several groups researching Variable Speed Control Moment Gyros which combine the gyroscopic effect of classical CMGs and the angular velocity effect of Reaction Wheels. Regardless of the momentum exchange device employed, the stability analysis of both Reaction Wheels and CMGs requires nonlinear stability theory when a fully agile system is desired. Both Reaction Wheels and CMGs suffer from saturation and gimbal lock problems which make controller design difficult at best.

While the dynamic analysis of momentum exchange devices is a challenge in itself, designing control laws for Reaction Wheels and CMGs is made that much more difficult due to the often time varying and nonlinear nature of the equations of motion. Therefore, it has been historically common to linearize the equations of motion around a single operating point, or to command the spacecraft in such a way as to decouple the angular velocities about orthogonal axes. Recently several groups have developed full state feedback laws which guarantee asymptotic stability. For example, Wie has shown asymptotic stability can be achieved with full state feedback when different conditions on the inertia matrix of the system are satisfied. To overcome gimbal lock and reaction wheel speed saturation, several approaches have been developed. See for example...

The adage, "a proof is worth a thousand simulations", remains very true in the world of control theory and controller design. If a stability proof can be formulated, all dynamic simulations should experimentally demonstrate stability properties. However, to gain confidence in a particular control law in a real world setting complete with noise, other disturbances and errors associated with measurements and control authority, groups typically abide by the accepted practice of building numeric and physical simulations.

The Cal Poly Spacecraft Dynamic Simulator, CP/SADS, has been designed to serve as a way of quickly assembling experimental data while following the three pronged approach outlined above. Ideally, theoretical analysis would precede numeric simulation followed by experimental data and finally, the spacecraft could be built. This paper outlines the design and development of such an experimental test bed. The kinematic and dynamic analysis used to predict the motion of the simulator is presented. There are significant differences in the dynamics of rigid bodies in space and rigid bodies on ground. The constant freefall of spacecraft in earth orbit is the most significant difference. Spacecraft simulators on the surface of the earth must be supported and hence a gravity torque is common. Our approach to this significant problem will be discussed in more detail. The design and development of the CP/SADS will also be presented. Each subsystem will be covered in detail, including all structural components, the reaction wheels, power supply, sensors and actuators, data acquisition, and real time control. The preliminary results and testing procedures developed to date will also be presented. While the CP/SADS is still under development, preliminary test results agree well with predicted performance. Finally, the future of the CP/SADS will be outlined. Our group hopes to develop a 'plug-and-play' approach based on using the CP/SADS for attitude control algorithm testing.

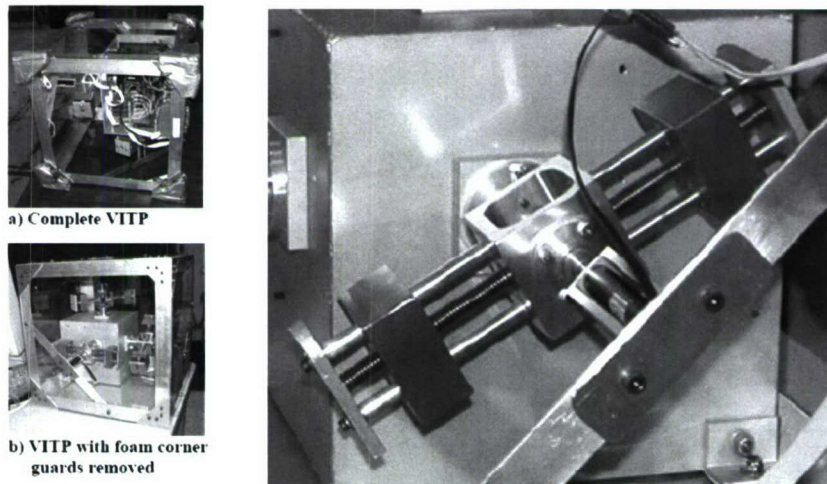
### **B. Past and Current Spacecraft Dynamic Simulators**

Reaction wheel control devices have been around since the Apollo missions in the 1960's. The four-wheel pyramidal reaction wheel configuration is a common type of reaction wheel system and is used in many spacecraft today. The advantages of a four-wheel pyramidal configuration for attitude control devices are its light weight and

redundancy since the pyramidal configuration has four motors rather than the three that are minimally needed for full attitude control.

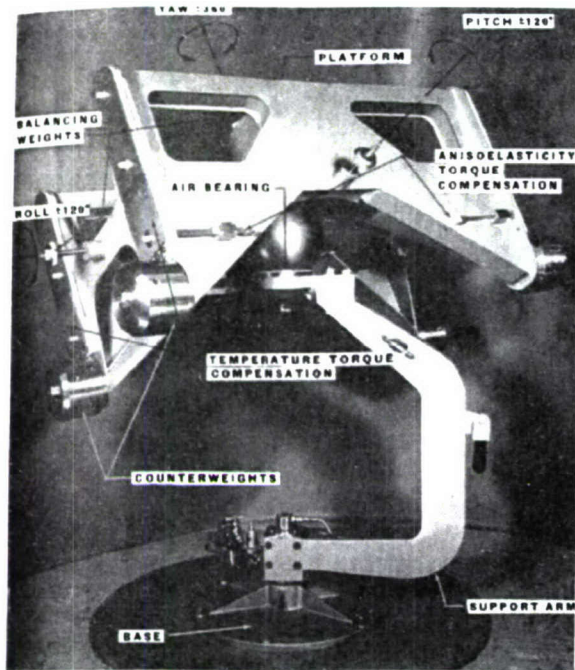
The Control Momentum Gyro (CMG) is a momentum exchange device that is closely related to the reaction wheel. Like the reaction wheel, the CMG was developed to make precise changes in attitude control of the spacecraft. The CMG differs from a reaction wheel in that the inertia wheel's spin is not fixed on a precise angle relative to the spacecraft. Instead the CMG's are nutated by a servomotor. The inertia wheels are rotated at constant angular velocity and the nutation creates a moment perpendicular to the spin and the nutation axes. CMG's have an advantage over Reaction Wheels since they are not limited by saturation of the motors. However, depending on the configuration CMG's can be plagued by gimbal lock, which occurs when two or more gimbal axes have been aligned. Reaction Wheels do not suffer from gimbal lock. Gimbal lock is usually combated by using multiple CMG's. Future versions of the Cal Poly Reaction Wheel Platform may incorporate CMG wheels.

New and different reaction wheel control algorithms and products have been developed in recent years. A recent variation of the reaction wheel is the Variable Inertia Reaction Wheel (VIRW) developed by researchers at the Georgia Institute of Technology (Figure 1<sup>2</sup>) as part of the Variable Inertia Test Platform (VITP). The VIRW has the ability to change the inertias of any of the wheels. It does this by way of two masses which are allowed to transverse the length of the wheel while it is spinning. This type of system requires balancing techniques that are complicated and difficult to implement. However, if the added balancing complexity is tolerable, the VIRW provides two distinct advantages: variable precision and vehicle reorientation without the change in angular speed of the wheel. Variable precision is critical in reducing jitter, which allows for fine pointing of the spacecraft.<sup>2</sup>



**Figure 1: Images of variable inertia test platform**

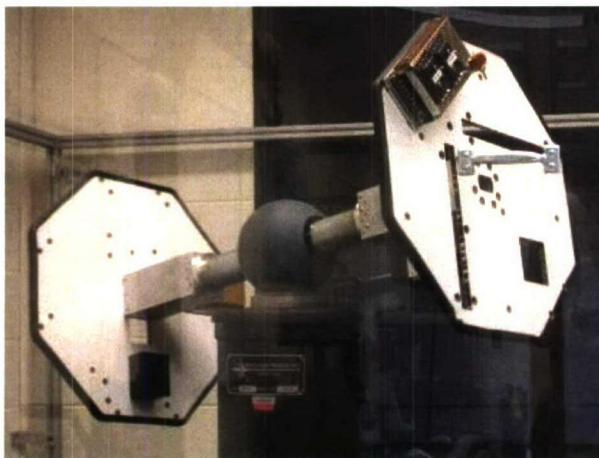
Regardless of what type of momentum exchange device is developed, to simulate the rotational motion associated with spacecraft in orbit, the momentum exchange device is typically mounted on some sort of bearing surface. The combination of a reaction wheel of any flavor or a CMG with two or three axis low friction bearing surface allows for spacecraft attitude dynamics simulations. Figure 2 is a photo of one of the first spacecraft dynamics simulator. The Satellite Motion Simulator developed by NASA in the 1960's was used as a test bed to design and develop satellite attitude control laws during the early days of the United States Space Program.



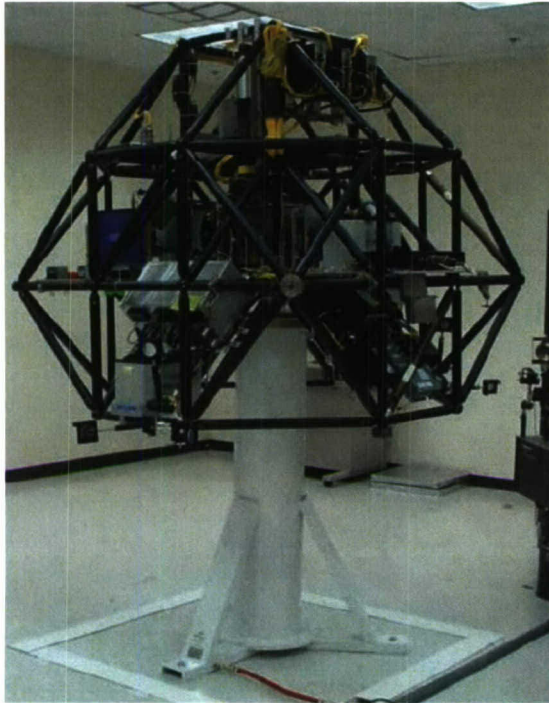
### **SATELLITE MOTION SIMULATOR**

**Figure 2: 1960 NASA Air Bearing Satellite Simulator**

Figure 3 is a recent photo taken from the Virginia Tech Space Systems Simulation Laboratory (SSSL) website<sup>10</sup>. The SSSL maintains several space dynamics simulators, the Distributed Spacecraft Attitude Control System Simulator (DSACS), Whorl-I and Whorl-II, pictured in Figure 3. DSACS uses a spherical air bearing to simulate the attitude dynamics of a spacecraft while Whorl-I and Whorl-II employ a dumbbell air bearing configuration to achieve 360 degrees of rotational motion in two axes. All three systems are used to simulate the attitude dynamics of spacecraft in orbit. Finally, Figure 4, is a photo of the Honeywell Space System Simulator currently under development. Notice the Honeywell simulator also has a modular design.



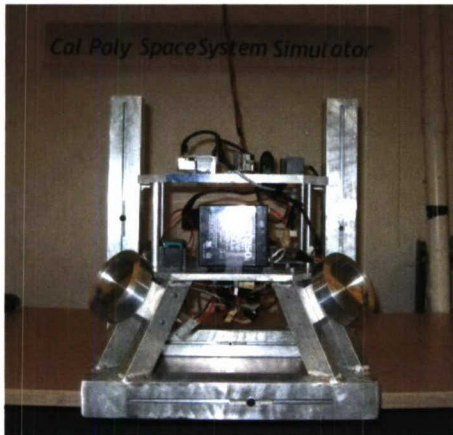
**Figure 3: Virginia Tech Dumbbell Style Simulator**



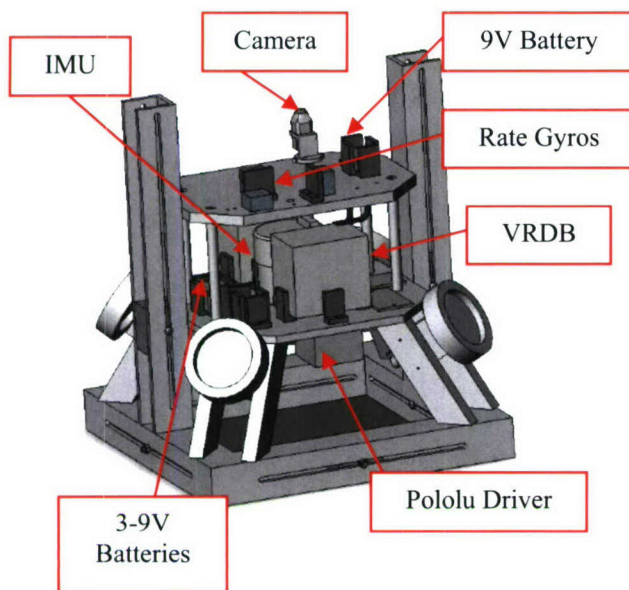
**Figure 4: Honeywell Space System Simulator, circa 2003**

### **C. An Overview of the Cal Poly Spacecraft Attitude Dynamics Simulator**

A recent photo of the CP/SADS is shown in Figure 5. A solid model of the CP/SADS is shown in Figure 6. Notice the CP/SADS structure rests on a spherical air bearing. Hence the CP/SADS has a full 360 degrees of motion about the z-axis and  $\pm 30$  degrees about both the x- and y-axes. The CP/SADS differs from the projects discussed above since it relies on a pyramidal reaction wheel configuration. The reaction wheels are placed along the edges of what amounts to a square pyramid. The pyramidal configuration of the reaction wheels and the associated structure is inherently symmetric about all three axes. Only the relatively small masses associated with electronic components and the power supply have to be balanced about the center of gravity of the air bearing. More discussion on balancing is found below. The CP/SADS also has a coarse balancing system built into the structure which is used to modify the center of gravity of the CP/SADS. Appendages can be added to the platform to test different satellite configurations such as solar arrays. More detail on all of the CP/SADS subsystems and design can be found in subsequent sections.



**Figure 5: The Cal Poly Spacecraft Attitude Dynamics Simulator**



**Figure 6: Solid Model of the Cal Poly Spacecraft Attitude Dynamics Simulator**

## **II. Spacecraft Equations of Motion: In Space vs. the Simulator**

### **A. Dynamics**

In this section the equations of motion of a satellite are developed. The Dynamic equations and the kinematic equations used to model the CP/SADS are presented. The significant differences in the equations of motion of a spacecraft in orbit and a spacecraft simulator on the ground are discussed in different sections. The approach taken here to model the dynamics of the simulator follows a fairly standard derivation of Euler's Equations for rigid body rotational dynamics. To model the kinematics of the simulator, a quaternion representation is developed.

#### *1. Reaction Wheel Dynamics*

The pyramidal reaction wheel design is favored over a three-axis reaction wheel design due to its inherent symmetry and increase in torque along each rotational axis. The symmetry and increased torque come with the price of added mass. Therefore, a final design will generally depend on the nature of the problem and any constraints applied to the system. A four axis reaction wheel design also adds redundancy in case one motor fails before the mission is complete.

When looking at the x and y axes as shown in Figure 7, all four wheels can be arranged at equal angles from each other ( $90^\circ$  apart) which allows for equal torque in the x and y planes. Since the platform is approximately symmetric in the x and y axes the principal moments of inertia in the x and y directions are assumed equal.

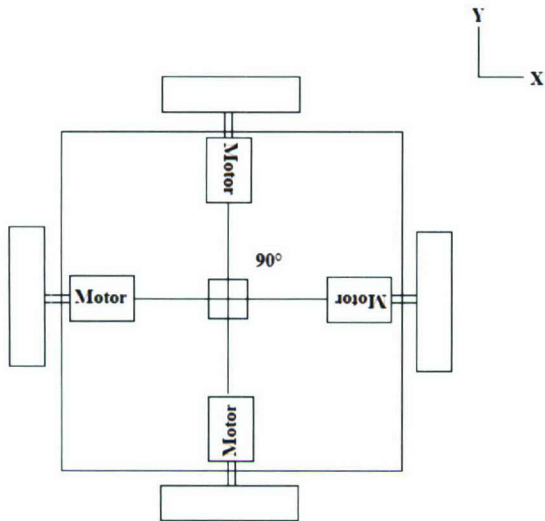


Figure 7: The four motors arranged 90° apart on the x and y axes

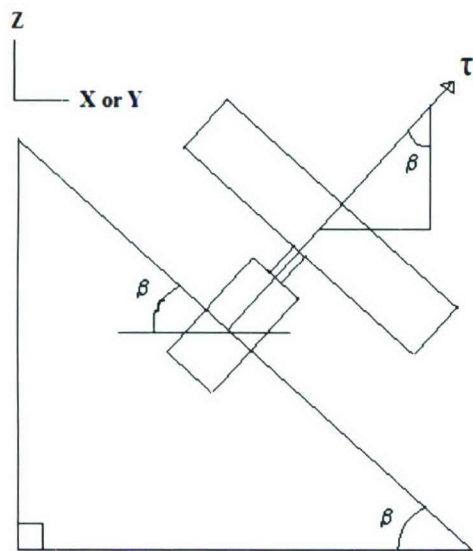


Figure 8: Motor tilted into the z axis by angle,  $\beta$

To apply a moment in only the z-axis all four, or any opposite two motors, are actuated simultaneously. The motor tilt angle  $\beta$  is formed as the angle from the motor spin axis to the z-axis. If the proper motor tilt angle is used, then the moment applied in the z-direction when all four motors accelerate is equal to the moment applied when any opposite two motors apply a moment about the axis connecting the two motors. Hence, due to symmetry, with the right choice of motor tilt angle, the moments applied in all three axes are equal.

Since the reaction wheel motors are tilted up toward the z-axis, all four motors contribute to the moment applied along the z-axis while only two motors are capable of applying a torque in either the x- or y- direction. When the applied moment from each motor is accounted for, the total moment in the x-, y-, and z-directions is expressed as,

$$\begin{aligned} T_x &= J(\dot{\omega}_{w_1} + \dot{\omega}_{w_3})\sin(\beta) = I_x \dot{\omega}_x \\ T_y &= J(\dot{\omega}_{w_2} + \dot{\omega}_{w_4})\sin(\beta) = I_y \dot{\omega}_y \\ T_z &= J(\dot{\omega}_{w_1} + \dot{\omega}_{w_2} + \dot{\omega}_{w_3} + \dot{\omega}_{w_4})\cos(\beta) = I_z \dot{\omega}_z \end{aligned} \quad (1)$$

where the inertia of each wheel is the same and given as,  $J$ . The angular acceleration of each wheel is given by  $\dot{\omega}_{w_j}$  and the angular acceleration of the platform in the x-, y- and z-axes are given by  $\dot{\omega}_x$ ,  $\dot{\omega}_y$  and  $\dot{\omega}_z$  respectively. The principal moments of inertia of the platform, including the inertia of each wheel about the platform center of gravity, are given by  $I_x$ ,  $I_y$  and  $I_z$ . The principal axes lie close to the body frame of reference demonstrated in Figures 7 and 8 due to the symmetry of the platform.

If the angular acceleration in the z-axis is equal to the angular accelerations in either the y or x axes then the reaction wheel set has equal control authority in all three axes. Assuming the angular acceleration is the same for each reaction wheel and setting the angular acceleration in the x- or y-direction equal to the angular acceleration in the z-direction as given in Eq. (1) leads to the relationship,

$$\frac{\sin(\beta)}{I_{x,y}} = \frac{2 \cos(\beta)}{I_z}. \quad (2)$$

Solving for the angle  $\alpha$  which insures equal control authority results in,

$$\alpha = \tan^{-1} \left( 2 \frac{I_{x,y}}{I_z} \right). \quad (3)$$

The ideal motor tilt angle is dependent on the relationship between the x- and y- principal inertias and the z- principal inertia. If  $I_{x,y} = I_z$  the angle,  $\beta$ , is  $63.4^\circ$ . However, if  $I_{x,y} < I_z$  then  $\beta$  is less and conversely if  $I_{x,y} > I_z$  then  $\beta$  is greater.

Initial design concepts of the CP/SADS indicated that the inertia about the z-axis would be greater than the inertias in the x- and y-axes. The inertias in the x- and y-axes were assumed equal. Therefore, a motor tilt angle of  $57.0^\circ$  was chosen for  $\alpha$  and written down as a hard requirement in the specifications list given to the structures team<sup>5</sup>. The final design advocated by the structures team had an inertia in the z-direction of  $0.5952 \text{ kg-m}^2$  and  $0.5528 \text{ kg-m}^2$  in the x and y directions. Hence, the ideal motor tilt angle is actually  $61.7^\circ$ . Since the ideal motor tilt angle is slightly greater than  $57.0^\circ$  the CP/SADS has slightly more control authority in the z-axis than in the x- and y-axes.

## 2. Actuators Dynamics

In order to achieve an angular acceleration of  $0.1 \text{ rad/sec}^2$  and an angular velocity of  $0.1 \text{ rad/sec}$  as given in the specifications list a detailed analysis was done to insure proper motor sizing. Using linear graphing methodology a state equation was developed to relate input voltage to output torque on the platform<sup>3</sup>. Figure 9 shows the linear graph of a motor with a gear train.

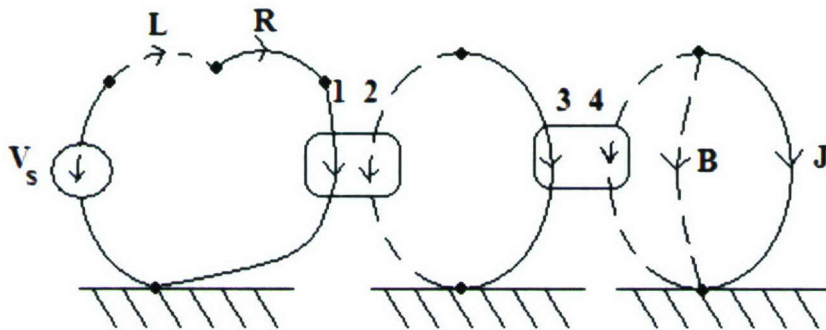


Figure 9: Linear graph of motor with gear train

The moment of inertia,  $J$ , is the combination of the reaction wheel inertia and motor shaft inertia. Viscous damping is given by  $B$ . The resistance of the motor windings is represented by  $R$  and inductance produced by the

current flowing around the winding is represented by  $L$ . The input voltage is given by  $V_s$ . The two transducers shown by the numbers 3-4 and 1-2 are the two-port elements which represent the gear train and the motor respectively. Since the transduction process relates across-variables to across-variables the connection between both 1-2 and 3-4 are related by two-port transformers defined by,

$$\begin{bmatrix} v_1 \\ f_1 \end{bmatrix} = \begin{bmatrix} TF & 0 \\ 0 & -1/TF \end{bmatrix} \begin{bmatrix} v_2 \\ f_2 \end{bmatrix} \quad (4)$$

Equation (4) relates the through-variables,  $f$ , to the across-variables,  $v$ . The transformer ratio is given by,  $TF$ , where  $TF$  equals the back-EMF constant,  $K_E$ , and the torque constant,  $K_M$ , for the motor. The two-port transformer relationship for the motor is then given as,

$$\begin{bmatrix} V_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} K_E & 0 \\ 0 & -1/K_M \end{bmatrix} \begin{bmatrix} \Omega_2 \\ T_2 \end{bmatrix}. \quad (5)$$

The transformer ratio also equals the gear ratio,  $GR$ , for the gear train. The two-port transformer relationship for the gear train is also given as,

$$\begin{bmatrix} \Omega_3 \\ T_3 \end{bmatrix} = \begin{bmatrix} GR & 0 \\ 0 & -1/GR \end{bmatrix} \begin{bmatrix} \Omega_4 \\ T_4 \end{bmatrix}. \quad (6)$$

From the linear graph, eight elemental equations describe the energy flow and are given as,

$$\begin{aligned} \frac{di}{dt} &= \frac{1}{L} V_L, & \frac{d\Omega_J}{dt} &= \frac{1}{J} T_J, \\ T_B &= B\Omega_B, & V_R &= Ri_R, \\ V_1 &= K_E \Omega_2, & T_2 &= -K_M i_1, \\ \Omega_3 &= GR\Omega_4, & T_4 &= -GRT_3, \end{aligned} \quad (7)$$

Along with the elemental equations we can write down five compatibility equations given as,

$$\begin{aligned} V_1 &= -V_R - V_L + V_s \\ \Omega_2 &= \Omega_3 \\ \Omega_4 &= \Omega_B \\ \Omega_B &= \Omega_J \\ \Omega_4 &= \Omega_J \end{aligned} \quad (8)$$

From the linear graph we also get five continuity equations given as,

$$\begin{aligned}
i_1 &= i_R \\
i_L &= i_R \\
i_2 &= i_3 \\
T_J &= -T_4 - T_B \\
T_3 &= -T_2
\end{aligned} \tag{9}$$

The combination of the elemental equations (7), compatibility equations (8) and the continuity equations (9) results in the two coupled differential equations,

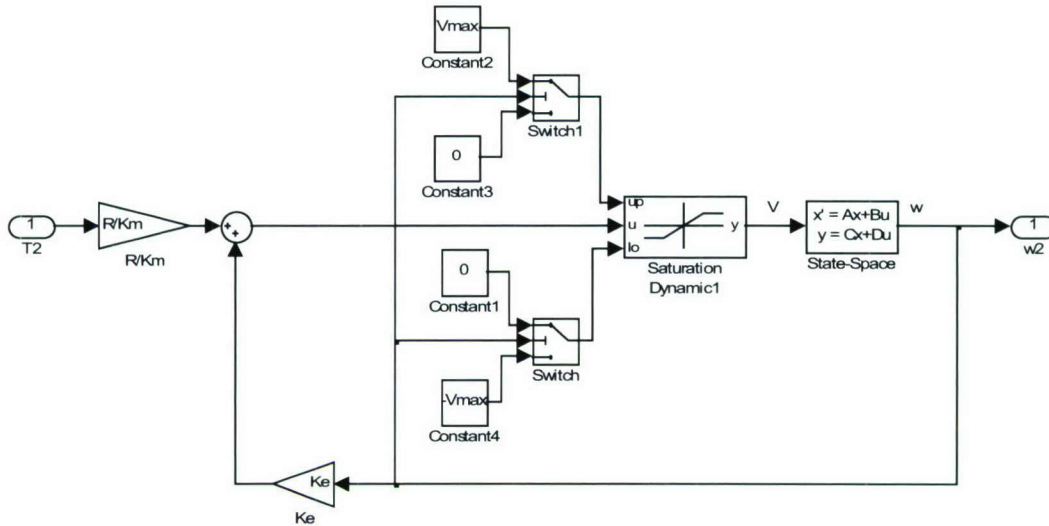
$$\begin{aligned}
\frac{di_L}{dt} &= \frac{1}{L}(-K_E \cdot GR \cdot \Omega_J - R \cdot i_L + V_S) \\
\frac{d\Omega_J}{dt} &= \frac{1}{J}(K_M \cdot GR \cdot i_L - B \cdot \Omega_J)
\end{aligned} \tag{10}$$

The two first order ordinary differential equations from Eq. (10) are written in state space form as,

$$\frac{d}{dt} \begin{bmatrix} i_L \\ \Omega_J \end{bmatrix} = \begin{bmatrix} \frac{-R}{L} & -GR \frac{K_E}{L} \\ GR \frac{K_M}{J} & \frac{-B}{J} \end{bmatrix} \begin{bmatrix} i_L \\ \Omega_J \end{bmatrix} + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} V_S \tag{11}$$

The angular velocity of each reaction wheel is found by integrating Eq. (11). Finally, the angular momentum of each reaction wheel is determined by multiplying the moment of inertia of each reaction wheel with the appropriate angular velocity. Since angular momentum is conserved between the reaction wheel and the platform the angular velocity of the platform is calculated after taking into account the number of reaction wheels and their orientation with respect to the platform's body frame.

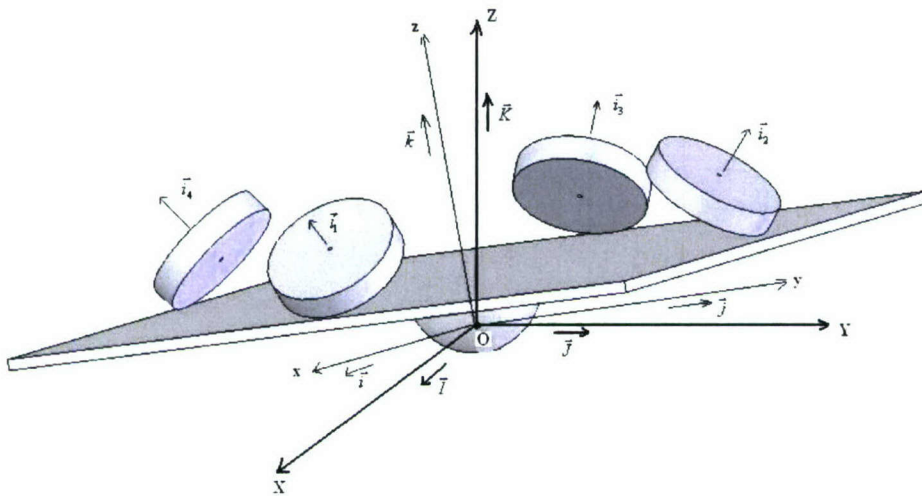
Figure 10 is a MatLab-Simulink block diagram which takes the input torque command and determines the angular velocity of each the reaction wheel. When a torque is commanded to the motor the code takes the torque and converts this torque into an equivalent input voltage will drive the motor to the commanded torque. Each motor accepts a voltage ranges from 0V to +24V when the motor is spinning in the positive direction and 0V to -24V when the motor is spinning in the negative direction. When the motor is not spinning the voltage range is from +24V to -24V. The two trigger switches change the dynamic saturation minimum and maximum depending on the speed of the motor. The voltage applied to the motor is not allowed to change polarity when the motor is spinning since such an applied voltage would create a large current spike and damage the insulation on the wire in the motor coil. The state-space block in Figure 5 represents Eq. (11) and transfers the voltage to an angular velocity of the reaction wheel.



**Figure 10: Motor simulation with reaction wheel attached**

### 3. CP/SADS Dynamics

To derive the dynamic equations of motion of the CP/SADS, two assumptions were made. First, the center of gravity of the CP/SADS structure was aligned with the pivot point of the spherical air bearing supporting the structure. Therefore, the gravity torque due to misalignments is zero. Second, the coordinate system shown in Figure 11 (adapted from Healy<sup>4</sup>) with basis vectors  $\vec{I}, \vec{J}, \vec{K}$  is an inertial reference frame. Next, the body fixed frame, B, was defined by the basis vectors  $\vec{i}, \vec{j}, \vec{k}$ . Both frames have fixed origins at the center of rotation of the air bearing, O. The unit vectors through the axis of rotation of the reaction wheels are  $\{\vec{i}_1, \vec{i}_2, \vec{i}_3, \vec{i}_4\}$ .



**Figure 11: Reference Frame Schematic**

Using the inertial and body frames the equations of a motion for a rigid body are derived from Euler's Equation where the applied torque must equal the time rate of change of the angular momentum in the inertial frame. The angular momentum of a rigid body denoted by  $\vec{h}$  is given as,

$$\vec{h} = I \cdot \vec{\omega} \quad (12)$$

In Eq. (12), the platform's inertia matrix,  $I$ , is not diagonal; it includes the products of inertia as well as the moments of inertia and is given by,

$$I = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy} & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz} \end{bmatrix}. \quad (13)$$

In the case of the CP/SADS the rigid body has four reaction wheels, therefore, Eq. (12) must be expanded to include the angular momentum of each reaction wheel. Hence,

$$\vec{h} = I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)}. \quad (14)$$

Equation (14) states that the angular momentum  $\vec{h}$  equals the product of the inertia matrix of the entire platform about the platform center of gravity,  $I_G$ , with the angular velocity of the platform,  $\vec{\omega}^{(p)}$ , added to the sum of each reaction wheel's angular momentum about the motor spin axis,  $J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)}$ .

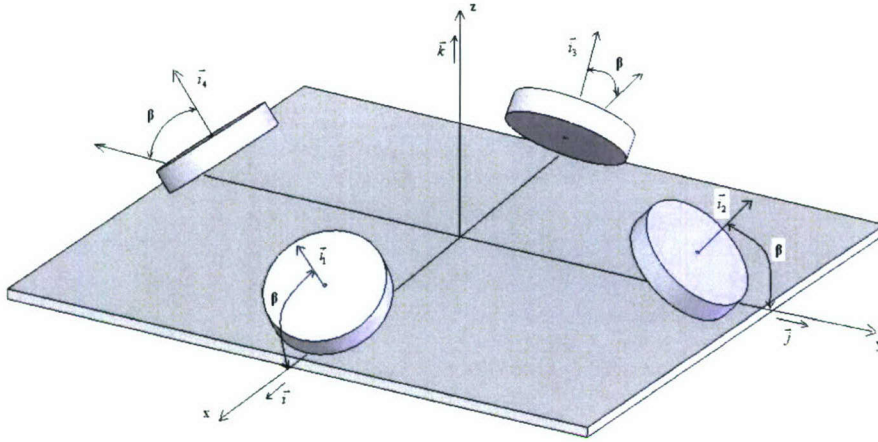
Taking the time derivative of the angular momentum in Eq. (14) gives the moment equation with respect to the inertial frame. By applying the proper derivative and cross product to account for the rotating body reference frame used to express the inertias, the dynamic equations of motion of the CP/SADS is expressed as,

$$\vec{M} = \frac{d}{dt} \left( I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)} \right) + \vec{\omega}^{(p)} \times \left( I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 J_{G_n}^{(n)} \cdot \vec{\omega}_{rel}^{(n)} \right). \quad (15)$$

Since all four reaction wheels on the platform are identical the moment of inertia of each wheel given by  $J_{G_n}^{(n)}$  can be written as  $J$ . Rewriting Eq. (15) to separate out the wheel dynamics from the platform dynamics yields,

$$\vec{M} = I_G \cdot \dot{\vec{\omega}}^{(p)} + \vec{\omega}^{(p)} \times I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 \left( J \cdot \dot{\vec{\omega}}_{rel}^{(n)} + \vec{\omega}^{(p)} \times J \cdot \vec{\omega}_{rel}^{(n)} \right). \quad (16)$$

The summation in Eq. (16) is valid when the reaction wheel and platform angular velocity vectors are expressed in the same coordinate system. Therefore, a coordinate transformation is performed to rotate the angular velocity of each reaction wheel expressed in the wheel frame to the body frame. Figure 12 shows the orientation of each wheel with respect to the platform<sup>4</sup>.



**Figure 12: Reaction Wheel Pyramid Formation**

In Figure 12, the motor tilt angle is again represented by  $\beta$  and is the angle between each reaction wheel's spin axis and the corresponding body axis it lies on. To determine the angular velocity of each reaction wheel in the body frame, the following rotation matrices were developed to rotate from the reaction wheel spin axis to the body axis. The subscript  $n$  denotes the number of each wheel shown in Figure 12<sup>4</sup>.

$$\begin{aligned}
 R_{W1} &= \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_{W2} &= \begin{bmatrix} 0 & 1 & 0 \\ \cos \beta & 0 & -\sin \beta \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_{W3} &= \begin{bmatrix} -\cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_{W4} &= \begin{bmatrix} 0 & 1 & 0 \\ -\cos \beta & 0 & \sin \beta \\ \sin \beta & 0 & \cos \beta \end{bmatrix}
 \end{aligned} \quad (17)$$

Since the reaction wheels are symmetric their inertia matrix is diagonal in the reaction wheel axis. The major principal axis is the same as the torque axis. The angular momentum equation for each wheel in its own coordinate frame is given by,

$$h_n = J \cdot \vec{\omega}_{rel}^{(n)} = \begin{bmatrix} J_{11} & 0 & 0 \\ 0 & J_{22} & 0 \\ 0 & 0 & J_{33} \end{bmatrix} \begin{bmatrix} \omega_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{11} \omega_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} \quad (18)$$

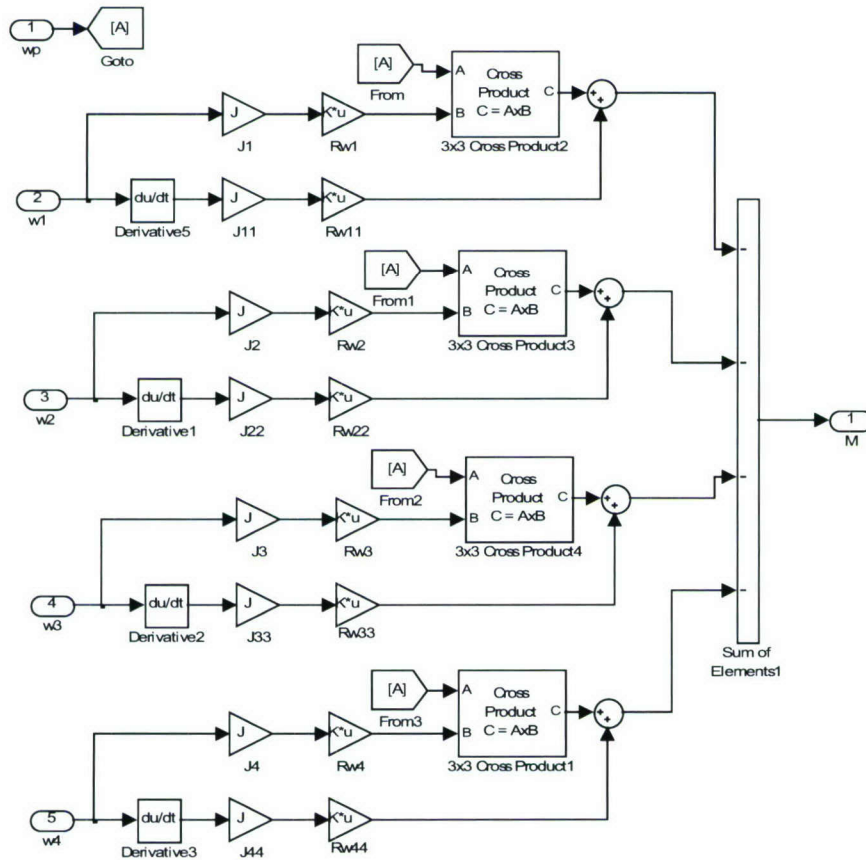
Taking the time derivative of the momentum equation gives the torque contribution of each reaction wheel.

$$T_n = J \cdot \dot{\vec{\omega}}_{rel}^{(n)} = \begin{bmatrix} J_{11} & 0 & 0 \\ 0 & J_{22} & 0 \\ 0 & 0 & J_{33} \end{bmatrix} \begin{bmatrix} \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{11} \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix}. \quad (19)$$

Finally, the equations of motion for the CP/SADS are found after substituting Eqs. (18) and (19), along with the coordinate rotations from Eq. (17) into Eq.(16).

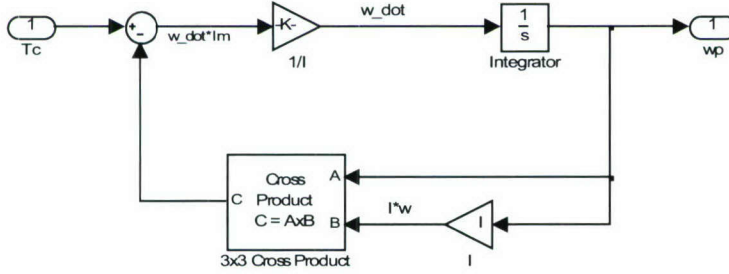
$$\vec{M} = I_G \cdot \dot{\vec{\omega}}^{(p)} + \vec{\omega}^{(p)} \times I_G \cdot \vec{\omega}^{(p)} + \sum_{n=1}^4 \left( R_{W_n} \cdot \begin{bmatrix} J_{11} \cdot \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} + \vec{\omega}^{(p)} \times R_{W_n} \cdot \begin{bmatrix} J_{11} \cdot \dot{\omega}_{rel}^{(n)} \\ 0 \\ 0 \end{bmatrix} \right). \quad (20)$$

Figure 13 is a Simulink block diagram which represents the reaction wheel dynamics given in the summation term on the right hand side of Eq. (20). The output of the subsystem represented in Figure 13 is the sum of moments for all four wheels with respect to the inertial frame.



**Figure 13: Reaction wheel dynamic simulation**

Figure 14 is the Simulink block diagram which models the dynamics of the platform and corresponds to the terms outside the summation on the right hand side of Eq. (20). In general the Simulink model allows for commanded torques from the reaction wheels and any external moment.



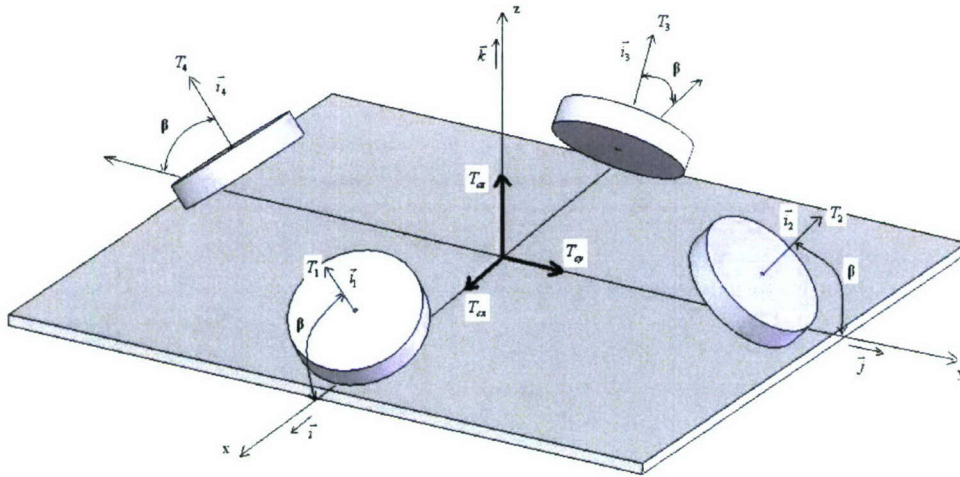
**Figure 14: Dynamic platform simulation**

#### 4. Commanded Torque Distribution and Minimization

Since the CP/SADS has four reaction wheels the commanded torque requirement should be distributed to each motor in such a way to minimize the energy used by the motors. With respect to the geometry of the platform in shown in Figure 15 the command torque on each body axis is,

$$\begin{aligned} T_{cx} &= T_1 \cos(\beta) - T_3 \cos(\beta) \\ T_{cy} &= T_2 \cos(\beta) - T_4 \cos(\beta) \\ T_{cz} &= T_1 \sin(\beta) + T_2 \sin(\beta) + T_3 \sin(\beta) + T_4 \sin(\beta) \end{aligned} \quad (21)$$

The torques,  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$  are the torques about each reaction wheel spin axis.



**Figure 15: Torque Relationships on a Reaction Wheel Pyramid**

In matrix form Eq. (21) is represented as,

$$\begin{bmatrix} \hat{T}_{cx} \\ \hat{T}_{cy} \\ \hat{T}_{cz} \end{bmatrix} = \begin{bmatrix} T_{cx} / \cos(\beta) \\ T_{cy} / \cos(\beta) \\ T_{cz} / \sin(\beta) \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = [A_w] T \quad (22)$$

Since the matrix  $A_w$  is not square it cannot be inverted. In order to find the motor torques of each motor an auxiliary equation is added. To minimize the required torque each motor applies, we can use the method of Calculus of Variations to minimize the norm of the torque vector,  $\vec{T}$ . Minimization of the norm of the vector  $\vec{T}$  is accomplished using the Lagrangian,

$$L = H + \lambda_1 g_1 + \lambda_2 g_2 + \lambda_3 g_3. \quad (23)$$

The Hamiltonian,  $H$ , from Eq. (23) is expressed as,

$$H = \sum_{n=1}^4 T_n^2 \quad (24)$$

Therefore, the Lagrange coefficients from Eq. (23) are derived from Eq. (22) and Eq. (24) as,

$$g_1 = T_1 - T_3 - \hat{T}_{cx}$$

$$g_2 = T_2 - T_4 - \hat{T}_{cy}$$

$$g_3 = T_1 + T_2 + T_3 + T_4 - \hat{T}_{cz}$$

The partial derivatives of the Lagrangian with respect to the torque on each motor generate the solution which satisfies the conditions for minimizing the Hamiltonian. The partial derivatives are,

$$\begin{aligned} \frac{\partial L}{\partial T_1} &= 2T_1 + \lambda_1 + \lambda_3 = 0 \\ \frac{\partial L}{\partial T_2} &= 2T_2 + \lambda_2 + \lambda_3 = 0 \\ \frac{\partial L}{\partial T_3} &= 2T_3 - \lambda_1 + \lambda_3 = 0 \\ \frac{\partial L}{\partial T_4} &= 2T_4 - \lambda_2 + \lambda_3 = 0 \end{aligned} \quad (25)$$

The solution to Eq. (25) gives the desired auxiliary equation so the state matrix from Eq. (22) can be inverted. The final auxiliary condition is,

$$T_1 - T_2 + T_3 - T_4 = 0.$$

Therefore, Eq. (22) is rewritten in matrix form and inverted to give the relationship between the desired torque command to the rigid body and the torque command to each reaction wheel motor. The final reaction wheel torque command is given as,

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \hat{T}_{cx} \\ \hat{T}_{cy} \\ \hat{T}_{cz} \\ 0 \end{bmatrix}.$$

With the command torque for each reaction wheel determined, the dynamics of the CP/SADS are determined, if the simulator were in space. However, there are significant torques present due to gravity which must be considered when the simulator is in the lab, on earth.

### 5. Ground Simulation

The derivation of the dynamics of the CP/SADS in the above section assumes the pivot point of the spherical air bearing holding the CP/SADS is collocated with the center of gravity of the CP/SADS. The CP/SADS is balanced when the center of gravity and the pivot point of the CP/SADS are collocated. While balancing the CP/SADS is a desired goal, the location of the center of gravity is not easy to find. Relying on System Identification techniques, and following the methods developed by Tanygin and Williams<sup>11</sup>, Healy<sup>4</sup> has shown that the center of gravity of the CP/SADS can be located, and hence, moved by some sort of balancing technique. However, to date, the System Identification technique has not been applied to the CP/SADS. The development of the balancing device and algorithm is ongoing.

To compensate for the unbalanced simulator the preliminary results discussed in subsequent sections only considers rotations about the z-axis of the CP/SADS simulator. The symmetric design of the SP/SADS guarantees that the CP/SADS is nearly balanced in the x- and y-axes. Therefore, the coarse balancing system, (discussed below), was used to move the center of gravity of the CP/SADS well below the pivot point of the air bearing. Hence, the CP/SADS was passively stable about the x- and y-axes. Essentially, the CP/SADS behaves much like a pendulum about the x- and y-axes.

For a full analysis and discussion of the gravity torque introduced into the equations of motion of the simulator due to the dislocation of the center of gravity and the pivot point, see the paper by Healy<sup>4</sup>.

### B. Kinematics: Quaternions

To represent the kinematics of the CP/SADS a quaternion representation was used. The quaternion representation does not suffer from singularities in the solution to the kinematic equations. Also, the quaternion representation yields to nonlinear full state feedback stability analysis as described by Wie<sup>1</sup>. If the Euler angles are required for further analysis, the quaternion representation can yield that information.

In his book, Wie shows that the angular velocities, or body rates, of a rigid body in space are related to the quaternion and the quaternion derivatives by,

$$\begin{aligned}\omega_1 &= 2(\dot{q}_1 q_4 + \dot{q}_2 q_3 - \dot{q}_3 q_2 - \dot{q}_4 q_1) \\ \omega_2 &= 2(\dot{q}_2 q_4 + \dot{q}_3 q_1 - \dot{q}_1 q_3 - \dot{q}_4 q_2) \\ \omega_3 &= 2(\dot{q}_3 q_4 + \dot{q}_1 q_2 - \dot{q}_2 q_1 - \dot{q}_4 q_3)\end{aligned}\quad (26)$$

Since a quaternion is only a rotation when it's norm is equal to one, the derivative of the norm must then equal zero and is given as,

$$0 = 2(\dot{q}_1 q_1 + \dot{q}_2 q_2 + \dot{q}_3 q_3 + \dot{q}_4 q_4). \quad (27)$$

Combining Eqs. (26) and (27) into a vector matrix product gives,

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} q_4 & q_3 & -q_2 & -q_1 \\ -q_3 & q_4 & q_1 & -q_2 \\ q_2 & -q_1 & q_4 & -q_3 \\ q_1 & q_2 & q_3 & q_4 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \\ \dot{q}_4 \end{bmatrix}. \quad (28)$$

Taking the inverse of the quaternion matrix in order to relate the quaternion rates to the angular velocities and separation the vector part of the quaternion from the scalar part yields the differential equations,

$$\begin{aligned}\dot{\vec{q}} &= \frac{1}{2}(\vec{q}_4 \vec{\omega}_p - \vec{\omega}_p \times \vec{q}) \\ \dot{q}_4 &= -\frac{1}{2} \vec{\omega}_p^T \vec{q}\end{aligned}\quad (29)$$

### III. The Cal Poly Spacecraft Attitude Dynamic Simulator Design

#### A. Overall System Design

The CP/SADS is composed of five different sub-systems which include: the structure of the platform, the reaction wheel actuators, the power buses, sensors, and the data acquisition system. Figure 16 shows the entire hardware schematic of the platform. The voltage regulation/data board (VRDB) acts as the main hub for all low power and data lines. A 9 V battery is used to power the camera and its transmitter. The 12 V bus shown in Figure 16 is for future reference once the LN-200 IMU (Inertia Measurement Unit) is operational. The LN-200 was kindly donated by Northrop Grumman. The components behind the transparent boxes are components which will be added to the CP/SADS during future development stages.

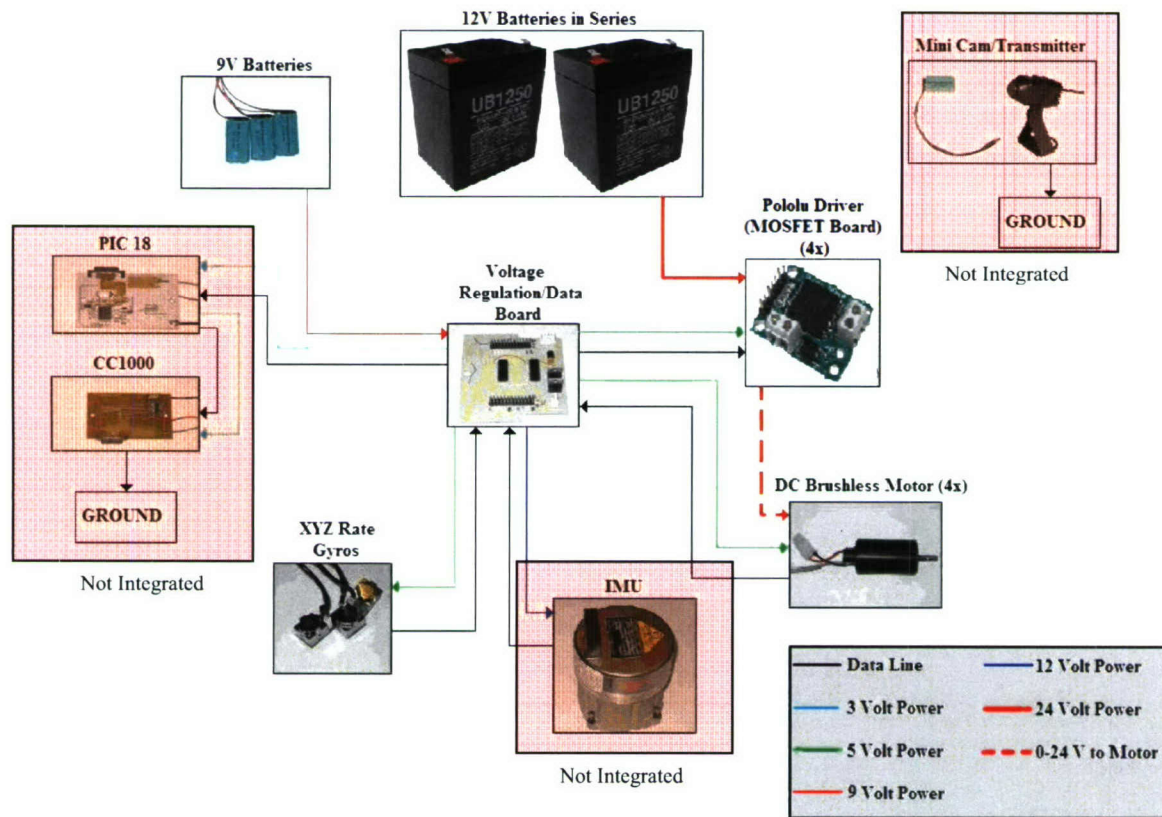


Figure 16: CP/SADS hardware schematic

#### B. Structural Design

The structure of the CP/SADS is used to hold and secure each sub-system; it is the body of the simulated spacecraft. Much of the design of the structure was determined by a specifications list. The specifications list was developed for the purpose of insuring hard requirements to the CP/SADS. Since most of the manufacturing and design was conducted by a mechanical engineering senior project team (see Martin, Quinto, and Nash<sup>5</sup>) a requirements document was established for clear communication between the Structure group and the System Integration group.

Many requirements were contemplated during the development of the specifications list; however, only five important requirements of the platform will be mentioned in this report. Important requirement #1 limited the weight to less than 10 Kg without including the adjusting weights, (discussed below). This requirement allowed for one user to pick up and operate the platform. The structure essentially acts as a simulated spacecraft structure and allows for mounting of other sensors, actuators, and other subsystems.

The structure was designed with angled supports which hold the reaction wheel motors at 57°. With the supports angled at 57° the reaction wheels would be angled at 33° from horizontal giving the motors some control authority in the z axis, see section II.A.1. The structure was also over built to limit resonant frequencies less than 10 Hz. Limiting any low vibration frequencies allowed the sensors to only measure the movement of the entire platform and not vibrations of the platform's structure which would induce error into the position and velocity measurements. Finally, the minimum acceleration of the CP/SADS about each axis should be 0.1 rad/sec<sup>2</sup> with a minimum velocity of 0.1 rad/sec upon startup. This requirement determined the inertia of the reaction wheels and was a factor in keeping the overall inertia of the platform down. The requirement was made to ensure substantial control authority over the platform.

### 1. *Passive Balancing*

The structure was designed with balancing in mind so it is symmetric about the x- and y-axes. Refer to Figure 6 for a solid model of the CP/SADS. Hence, a requirement was imposed on the Structural Design team to insure the CP/SADSs center of mass could be adjusted in case components were shifted. The Structural Design team came up with two channels on each axis. Each channel contains a weight capable of moving along the channel. Therefore, the center of gravity of the CP/SADS can be adjusted. The center of gravity adjustment system is designed to make coarse adjustments to the center of gravity of the CP/SADS. The weights were made sufficiently large to apply a moment of 2 N-m.

### 2. *Active Balancing*

While an active balancing system has not yet been fully developed, the structure of the CP/SADS was designed to accommodate the addition of such a system. An estimate of the mass of such a system was included in the solid modeling of the structural system and hence a location of the active balancing system was determined. The active balancing system of the CP/SADS is currently under development.

## C. **Actuator Design (Reaction Wheels)**

Four reaction wheels were designed to actuate the CP/SADS. For a complete analysis of the reaction wheel dynamics, see section II.A.1. This section outlines the sizing of the reaction wheels. In other words, the size and mass of the fly wheels was determined based on the imposed performance requirements.

Four Faulhaber 3863-24C DC Micromotors are used to actuate the platform. The motors are connected to the aluminum reaction wheels with inertias of 5.3552 Kg-cm<sup>2</sup>. The wheels trade off angular momentum by converting angular momentum of the wheels to angular momentum of the platform. The Faulhaber DC micromotors were used because of the high stall torque of 1.250 N-m. This is a substantial amount of torque for a 400 g motor and three times more than the equivalent servo motor. The stall torque is considerably more important for reaction wheel motors than the continuous torque, which is usually how motors are rated, because there is no external load applied to reaction wheel motors. The free spinning reaction wheels only produce a torque on the platform when the wheels are being accelerated and how fast the wheels can be accelerated is determined by the stall torque of the motors.

A large stall torque results in a large current draw from the motor. The stall torque is related to the applied current and the angular velocity of the rotor as,

$$T = K_M i - B\omega$$

where  $K_M$  is the motor torque constant, and  $B$  is the viscous friction of the motor. Here we assume the viscous friction of the motor is small and can be neglected.

The torque is the result of the large magnetic field produced by the current traveling around the motor windings. The motor windings naturally have a small amount of resistance. For steady state conditions the applied voltage of a DC motor is equal to the counter electromagnetic force, otherwise known as back EMF, which results from the spinning of the motor added to the resistance times the current, or,

$$V = K_E \omega + Ri$$

The resistivity of the Faulhaber motor is  $0.62 \Omega$ . Since the maximum voltage is 24 V the start up current of the motor can be as high as 38.71 A. This is an enormous amount of current but this only occurs for a brief moment as the motor catches speed. It is expected that the motors will run on average not more than 10 A.

In order to achieve a minimal platform angular acceleration of  $0.1 \text{ rad/sec}^2$  a large motor stall torque was needed. From Eq. (1) the CP/SADSs acceleration is determined by the motor stall torque and the overall platform's inertia. The motor stall torque can insure the platform's acceleration for a short time but the inertia of the reaction wheels are needed to determine how long this acceleration lasts and overall momentum transferred to the structure. By adding inertia to the reaction wheel it is possible to achieve the minimal platform angular velocity of  $0.1 \text{ rad/sec}$ . The more inertia the wheels have the more momentum each wheel can achieve. Likewise an important factor in wheel momentum is the maximum velocity of the wheel, which in the case of the Faulhaber 3863-24C DC Micromotors is 6700 rpm. This is also known as the threshold speed because the motor reaches saturation at this speed and is no longer capable of supplying torque in one of its two directions.

It is important to note that the motor has two forms of braking. The motor can be braked by allowing the motor to coast so that the motor loses momentum by the friction in the bearings and the air friction due to the reaction wheel's spin. The second form of braking is called active braking or shunting and is done by allowing the two terminals of the motor to be left either both high or both low. This causes the motor to act as a generator and back EMF is sent through the Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) board and back to the motor. The MOSFET's resistance added to the motor winding resistance act to stop the motor's spin. Another important note is that the time it takes the motor to go from an initial speed to a final speed is approximately equivalent to the time it takes to go from the final speed and brake back to the initial speed. This means that the acceleration and deceleration of the motor have equivalent time constants.

Attached to the motor is an encoder. The IE2-64 encoder is a magnetic encoder with a resolution of 64 lines per revolution. The encoder gives frequency modulated pulses which are interpreted as the motor's speed.

#### D. Power Subsystem

There are four different power buses on the CP/SADS: 3 V, 5 V, 9 V, 24 V. All power comes from batteries onboard the CP/SADS and is distributed on the voltage regulation/data board (VRDB). There are 6 batteries in all: two 12 V and four 9 V. The 12 V batteries supply power to the motors while the four 9 V batteries provide power to the electronics.

The 12 V Power Sonic, 5.0 A-h, sealed lead acid batteries are connected in series to make the 24 V line to the MOSFET board. A lead acid battery was chosen over other batteries due to its large discharge rate at 10 times its capacity or 10 c, which is 50 A. This is much greater than Nickel Metal Hydride (Ni-MH) batteries or Lithium Ion (Li-I) batteries which have discharge rates of around 3 c and 0.5 c respectively. Though lead acid batteries tend to weigh more per unit energy, Ah\*V, they still lead the way in discharge rate per unit mass due to their low internal resistance. Since total amount of energy in the battery is not as much of a concern as the amount of discharge current of the battery, the small lead acid battery is the obvious choice. The 50 A limit of the batteries is greater than the 40 A needed to run all four motors continuously.

The 24 V power bus is connected to four Pololu Motor Driver (MOSFET) Boards. Each Pololu driver acts as a voltage regulator for each motor. The Pololu driver has two n-fets and two p-fets which are pulsed to allow current to flow in both directions with minimal voltage loss.

The MOSFET board is controlled with two 5 V lines, IN 1 and IN 2 and one pulse width modulated 5 V line, EN (Enable). The control table, Table 1, shows the Boolean input for each line and its corresponding action.

Table 1: Motor Control Table

Enable	IN1	IN2	Action
H	L	L	Stop
H	L	H	Forward
H	H	L	Reverse
H	H	H	Stop
L	X	X	OFF

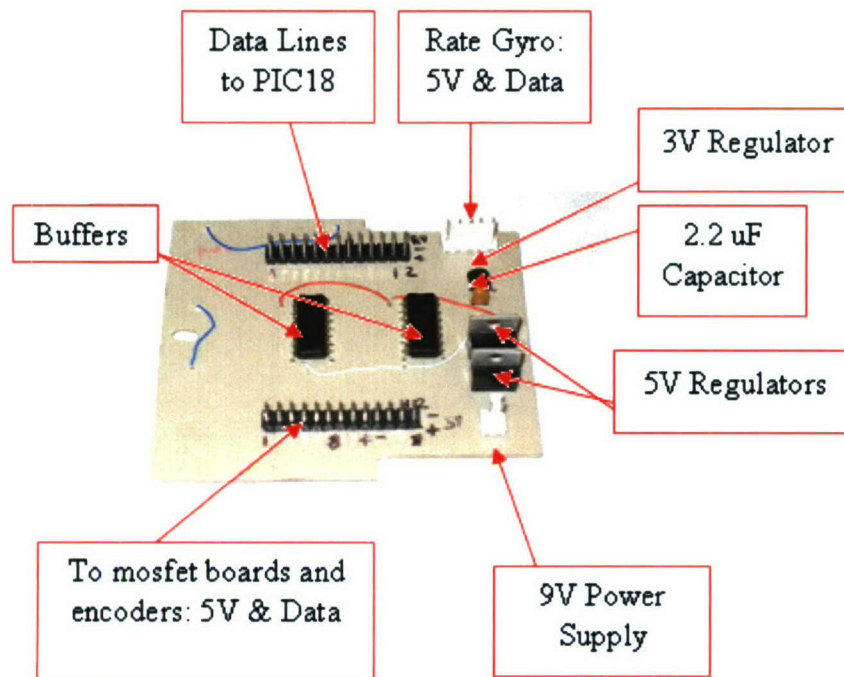
The VRDB is used as a central hub for all power and data lines with the exception of the 24 V motor power bus. There are three Ni-MH 9 V, 250 mAh batteries that power the VRDB.

The 9 V batteries reduce to 5 V and 3 V buses through the use of voltage regulators on the VRDB. The 5 V bus powers the MOSFET boards (Pololu Drivers), rate gyros, buffers on the VRDB, and the encoders. The power allocation of each component on the CP/SADS is given in Table 2. A total of 377 mA is needed to run all components on the CP/SADS. All three batteries have a total energy of 750 mAh, which puts each battery well within the maximum drain rate of 1 c, 250 mA.

Table 2: Power consumption of each component

Component	Amperage (mA)	Number	Total Amperage (mA)
D200	80	4	320
Buffer	10	2	20
Rate Gyro	7	3	21
Encoder	4	4	16
Total			377

Figure 17 shows the VRDB and the various components and connectors. The VRDB was a custom made board etched using copper solvent solution. On board the VRDB is a pair of buffers that act to protect the Programmable Interrupt Controller (PIC), which is currently not incorporated into the PRWP, as well as increase the voltage of the output PWM and I/O lines from 3 V to 5 V. The Texas Instrument CD4050BE Buffers each have 6 Operational Amplifiers (Op-Amps). Together the pair of buffers has enough terminals for all 12 output ports. A 2.2 uF capacitor is used to help steady the cyclic response of the 3 V regulator as well as absorb any transient spikes in the line. Capacitors were also added to the 5 V line for absorption purposes.



**Figure 17: Voltage Regulation/Data Board (VRDB)**

#### **E. Sensors**

Attitude determination of the CP/SADS is accomplished by three position and velocity sensors on board. The rate gyros and IMU are both dead reckoning sensors which determine position and velocity by forces that act on them. The third sensor, the star tracking camera, determines position and velocity by means of a reference point such as a star in a fixed reference frame. Over time any dead reckoning sensor will accumulate errors. As these errors grow the actual position and the sensed positions diverge. This phenomenon is known as drift. A star tracker is often used to update the known attitude of the spacecraft periodically. Attitude maintenance or upkeep is usually done on a half-day basis depending on the satellite.

Three Silicon Sensing CRS03-01 Rate Gyro sensors are placed on each of the three body axes to record angular rate measurements of the rotating platform. The rate gyros used are Micro-Electrical to Mechanical System (MEMS) devices. The rate gyros work using a Silicon Vibrating Structure Gyroscope (Si-VSG) which measures vibrations of an etched silicon ring in a magnetic field. Current is passed through the conductors on the ring to keep the ring vibrating at a constant 14 kHz.<sup>6</sup> This closed loop system produces a Rateout Voltage ( $V_o$ ) signal that is proportional to the angular rate being sensed. Determining angular position about each axis can be accomplished by integrating the voltage signal with respect to time. This process for determining angular position is less accurate over time as errors in the signal accumulate.

When integrated, the IMU can be used in conjunction with the rate gyros to give a more accurate position as time increases. The LN-200 has a drift rate of one degree per hour. The low drift rate of the LN-200 is due to the ring-laser gyroscope which measures the interference patterns of light emitted in opposite directions. The main cause of drift is due to thermal expansion problems which change the length the paths of light travel. The advantage over MEMS devices is the lack of moving parts which means no need for calibration other than the need for a zero-point adjustment.<sup>7</sup>

For a star tracker the ZTV Micro-Camera was chosen mostly for its small size, light weight, and refresh rate of 30 Hz which is fast enough for its application. The purpose of the camera is to locate the position of an LED placed above the platform. By tracking the LED's motion the platform's relative position can be determined relative to the inertial frame of the LED. With the camera attitude maintenance of the platform can be scheduled periodically to alleviate the error buildup of the rate sensors (the ZTV camera is not currently integrated)

#### **F. Data Acquisition**

One of the main objectives for future versions of the CP/SADS is to remove the need for an umbilical connection. In order to achieve this goal all data needs to be received and transmitted wirelessly from the platform

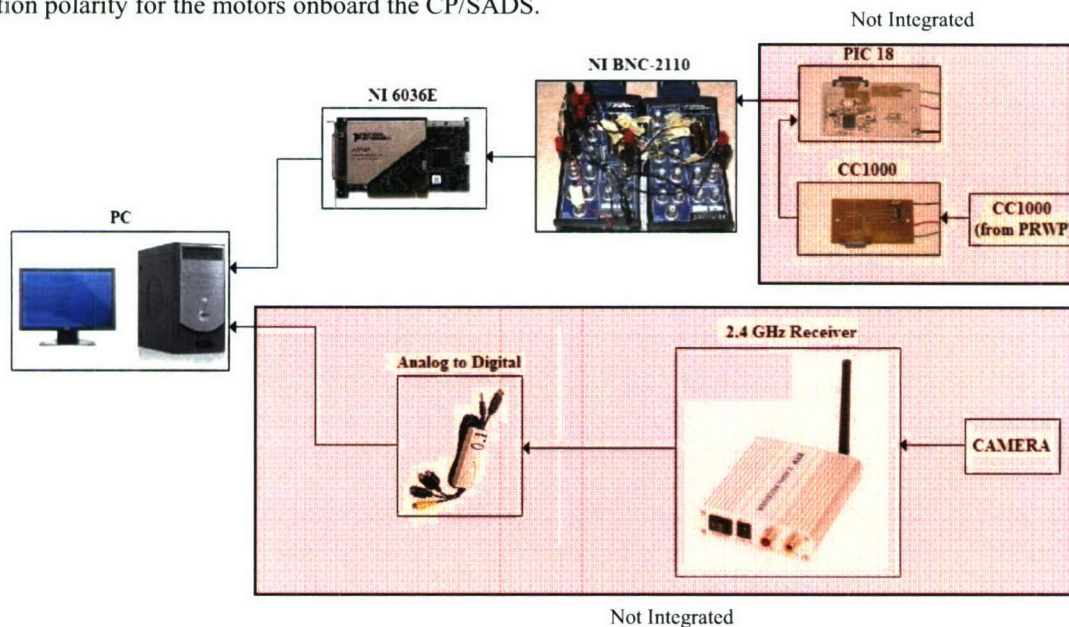
to the ground station and vice versa. Two transmitters will be used to send data from the platform to the ground. The first transmitter, which has already been installed on the platform, is embedded in the camera and it sends data to the ZTV 2.4 GHz Receiver positioned on the ground.

The ground based 2.4 GHz Receiver supplies an analog RCA output. The RCA output is converted into a digital signal by the Kworld analog-to-digital adapter that can be connected via a USB port to the PC. Using Matlab's image acquisition and processing toolboxes it is possible to process the data and find the position of the platform with respect to the inertial frame.

The second transmitter, which will be installed in a future version, is the Chipcon CC1000 which acts as a receiver and transmitter, otherwise known as transceiver. The CC1000 lies on a Prototype Circuit Board (PCB) that also holds the PIC 18. Data is captured by the PIC's A/D, digital I/O, and counter ports. The data is then sent to the transceiver where it is stored. The stored data is sent in packets to another ground based CC1000 transceiver (see **Error! Reference source not found.**). Likewise data can be sent from the ground to the platform if the process is reversed (For more information on the design and software of the CP2 Transceiver see Huerta<sup>8</sup>).

Currently the umbilical connection from the platform ties into a pair of National Instruments BNC-2110 connector boxes. The boxes individually hold two 16-bit analog outputs, sixteen 16-bit analog inputs, eight digital I/O lines and two 24-bit counters. The BNC-2110s connect to two PCI 6036E cards installed in the computer. The PCI cards are also made by National Instruments and work to interface with Matlab using Matlab's Real-Time Windows Target.

Figure 18 shows the proposed ground based hardware schematic. Data is received through both the ZTV 2.4 GHz receiver and the CC1000 transceiver and sent to the PC (The current version of the CP/SADS does not incorporate the CC1000 transceiver due to the many intricacies of the communication protocol; instead an umbilical cord is attached that connects the VRDB to the National Instruments BNC-2110 breakout board). Matlab works both to process the data and to close the loop by commanding the appropriate shunt and drive voltages as well as the direction polarity for the motors onboard the CP/SADS.



**Figure 18: Ground based hardware schematic**

### G. Full State Feedback Controller Design

Since the kinematics and dynamics of the platform are nonlinear it was important to first linearize the system to derive a control law using full-state feedback. Asymptotic stability of the nonlinear system is shown later using a Lyapunov function. The equations of motion of the platform are given by Eqs. (20) and (29). Linearizing the vectors  $\vec{q}$ ,  $q_4$ , and  $\vec{\omega}$  about  $q_4 = 1$  gives,

$$\vec{\omega}_p \approx \delta \vec{\omega}_p$$

$$\vec{q} \approx \delta \vec{q}$$

$$q_4 \approx 1 + \delta q_4$$

which results in the linearized equations of motion,

$$\begin{aligned} \vec{M} &\approx I_M \delta \dot{\vec{\omega}}_p = T_C + \vec{M}_{Ext} \\ \delta \dot{\vec{q}} &\approx \frac{1}{2} (\vec{\omega}_p) \\ \delta \ddot{q}_4 &\approx 0 \end{aligned} \quad (30)$$

Solving for  $\vec{\omega}_p$  and  $\dot{\vec{\omega}}_p$  results in,

$$\begin{aligned} \delta \vec{\omega}_p &= 2 \delta \dot{\vec{q}} \\ \delta \dot{\vec{\omega}}_p &= 2 \delta \ddot{\vec{q}} \end{aligned}$$

Applying the control torque,

$$\vec{T}_c = -K \vec{q}_e - C \vec{\omega}, (31)$$

and assuming no external moment is applied to the platform results in,

$$I_m \delta \ddot{\vec{\omega}} + C \delta \dot{\vec{\omega}} + K \delta \vec{q}_e = 0.$$

Substituting Eq. (3) into the dynamic equation gives,

$$\frac{d^2 \delta \vec{q}_e}{dt^2} + C I_M^{-1} \frac{d \delta \vec{q}_e}{dt} + \frac{1}{2} K I_m^{-1} \delta \vec{q}_e = 0. (32)$$

Relating Eq. (32) to the damping,  $\zeta$ , and natural frequency of the system,  $\omega_n$ , gives the control matrix constants K and C,

$$\begin{aligned} K &= 2 I_m \omega_n^2 = \begin{bmatrix} 2 I_1 \omega_n^2 & 0 & 0 \\ 0 & 2 I_2 \omega_n^2 & 0 \\ 0 & 0 & 2 I_3 \omega_n^2 \end{bmatrix} \\ C &= 2 I_m \zeta \omega_n = \begin{bmatrix} 2 I_1 \zeta \omega_n & 0 & 0 \\ 0 & 2 I_2 \zeta \omega_n & 0 \\ 0 & 0 & 2 I_3 \zeta \omega_n \end{bmatrix} \end{aligned} \quad (33)$$

The inertia matrix,  $I_m$ , of the platform is approximately a diagonal matrix due to the symmetry of the platform. For this reason only the moments of inertia were taken and all products of inertia were omitted. However, movement of the counter weights on the platform can change this symmetry. For cases when the platform is not

especially symmetric the principal moments of inertia can be found and a transformation matrix can be used to transform the control torques to the body axes.

A full picture of the simulation can be seen in Figure 18. Each subsystem is represented by equations derived earlier in this report. A gravity torque can be introduced but was not since the simulation was only taken about the z body axis due to the difficulty of making the platform neutrally stable with the large counter weights. A normalizing block was put into the quaternion loop because as the simulation progresses the quaternions tend to drift.

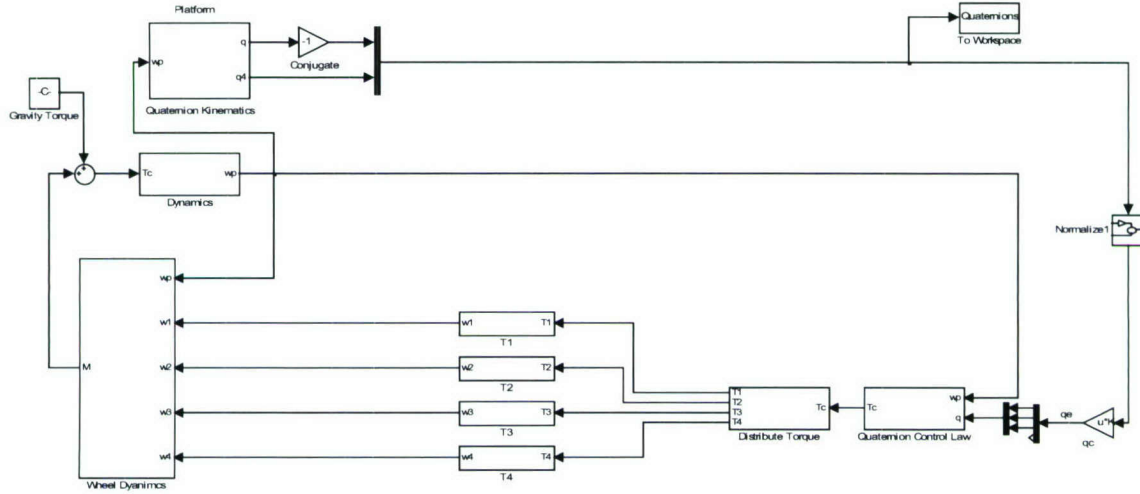


Figure 19: Control simulation of PRWP

#### H. Asymptotic Stability

From Wie<sup>1</sup>, asymptotic stability of the platform over a wider range of motion can be shown with the control law from Eq. (31) with gains defined from Eq. (33). First, generate a Lyapunov function of the form,

$$V(x) = \frac{I_1 \omega_1^2}{2k_1} + \frac{I_2 \omega_2^2}{2k_2} + \frac{I_3 \omega_3^2}{2k_3} + q_1^2 + q_2^2 + q_3^2 + (q_4 - 1)^2.$$

As shown in Wie<sup>1</sup>, a asymptotic stability is achieved for gains chosen such that,

$$\frac{I_2 - I_3}{2k_1} + \frac{I_3 - I_1}{2k_2} + \frac{I_1 - I_2}{2k_3} = 0$$

where  $k_1, k_2$ , and  $k_3$  are the diagonal terms in the K matrix from Eq. (33). The gains for the C matrix from Eq. (33) must simply be positive constants.

To show asymptotic stability the following criteria given by the Lyapunov-Barbalat theorem have to be satisfied along all trajectories,<sup>9</sup>

$$V(x) \geq 0$$

$$\dot{V}(x) \leq 0$$

$$\ddot{V}(x) \text{ is bounded}$$

The first criterion is satisfied by the Lyapunov function because the inertia matrices, control gains, and quaternions are all positive. Taking the derivative of the Lyapunov along trajectories function gives,

$$\dot{V}(x) = -\vec{\omega}_p^T C K^{-1} \vec{\omega}_p \geq 0$$

The second criterion is satisfied because the control gains  $C$ , and the inverse of  $K$  are positive definite matrices. The angular velocities of the platform are squared so this term is also positive. The second derivative of the Lyapunov function is given by

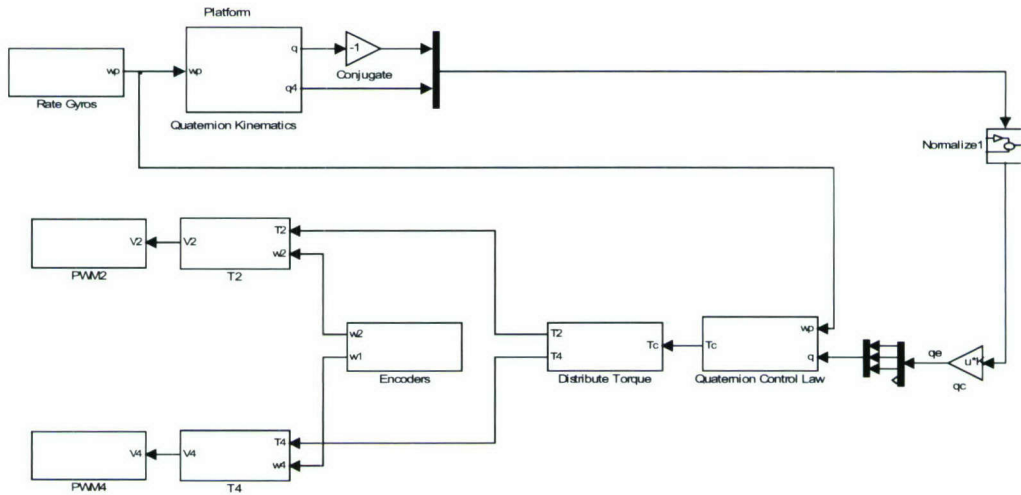
$$\ddot{V}(x) = -2\vec{\omega}_p^T C K^{-1} \vec{\omega}_p.$$

Barbalat's Lemma (criterion #3) is used to show asymptotic stability. The second derivative of the Lyapunov function is bounded meaning the Lyapunov derivative diminishes to zero as time approaches infinity, i.e.,  $\dot{V}(x) = 0$ . This means the angular velocity of the platform diminishes to zero as time approaches infinity,

$\vec{\omega}_p = 0$ . Therefore, the equations of motion of the platform in conjunction with the control law from Eq. (31) are asymptotically stable.

### I. MatLab Real Time Control Toolbox

To test the CP/SADS a control command was designed to move the platform through an angle of  $-90^\circ$  about the Z-axis. For this maneuver only two reaction wheels were used because the encoder on one of the motors was damaged during construction of the reaction wheel flywheels. For this reason the real time controller shown in Figure 20 differs from the simulated controller shown in Figure 19 in that there are only two motors (number 2 and number 4) instead of 4.



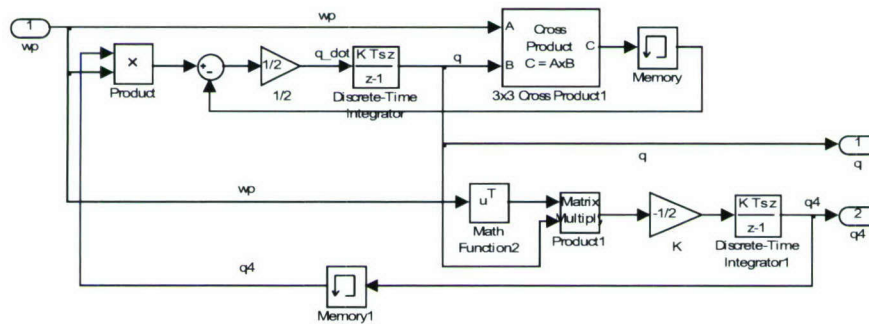
**Figure 20: Real Time PRWP Controller**

The real time controller is different from the simulated system since the real time system is based on digital data acquisition and command. The motor controllers use pulse width modulation to control motor speed. The rate gyros from Silicon Sensing were used to acquire body rates. The manufacturer specified that the angular rate ( $Ra$ ) is given by,

$$Ra = \frac{-\frac{1}{2} \times V_{dd} + V_O}{\left(SF \times \frac{V_{dd}}{5}\right)} \text{deg/sec.} \quad (34)$$

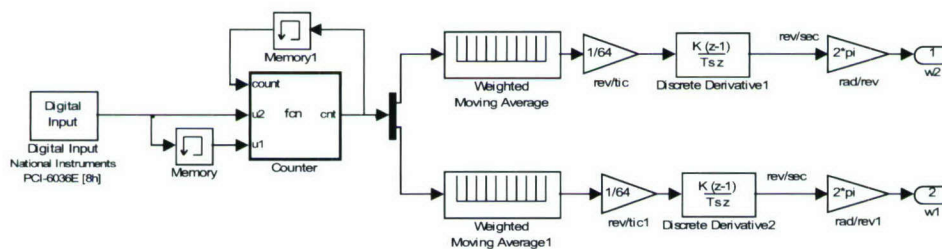
The voltages  $V_0$  and  $V_{dd}$  represent the output voltage of the rate gyro and the reference voltage supplied to the rate gyro respectively. The scale factor (SF) relates the difference in voltage to the angular rate in deg/sec. Figure 21 shows the rate gyro subsystem. The analog inputs represent the z axis gyro and the reference voltage. From Equ. (34), the voltages are converted into a z axis rate in deg/sec. This rate is then converted into rad/sec. A bias is added to the z axis rate to give a more accurate measurement. The bias is modified over time as the rate gyro heats up, but not during a command to the CP/SADS. To determine the correct bias the platform is put on a table and held stationary. A ten second sample is taken and integrated. A bias is then inserted to try to minimize the drift rate to zero. Without a bias the platform may drift as much as 0.3 deg/sec. Upon insertion of a bias the drift rate can be as low as 0.05 deg/sec (It is good to wait about twenty minutes to allow the gyro to come to thermal equilibrium before inserting a bias). The X and Y axis rates are set to zero and incorporated as one body rate. The rates are reversed because of the sensor orientation.

The quaternion kinematic subsystem shown in Figure 21 has only two changes from the quaternion subsystem from the simulated system. Two memory blocks are inserted into the feedback because Matlab is unable to perform an algebraic loop without a delay. In the simulated controller the model was running in continuous time but with the real time controller all models are running in discrete time.



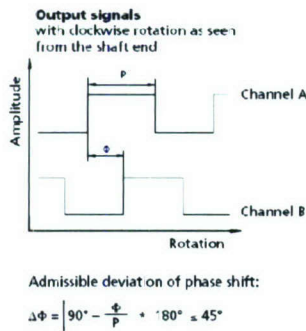
**Figure 21: Quaternion Kinematics Subsystem**

Figure 22 shows the encoder subsystem which converts each motor's encoder signals into the corresponding velocity of each motor in rad/sec. The counter embedded function looks for the rising edge of a pulse and then increments the count or decrements the count if the motor is reversed.



**Figure 22: Encoders Subsystem**

Since there are two encoders per motor (see Figure 23) direction is determined by relating channel A to channel B. Positive rotation according to Figure 11 is the opposite of that shown in Figure 23 so the output is reversed in the embedded code.



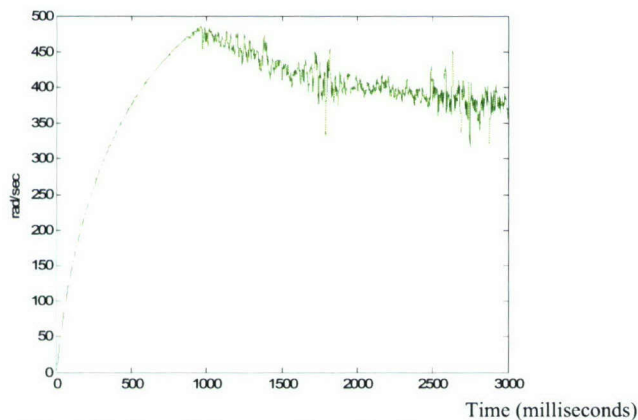
**Figure 23: Encoder output**

In the encoder subsystem the counts for each motor are separated and then a running average is taken to both smooth out the signal and delay the signal from 20 kHz to 100Hz. The counts are then divided by 64 because there are 64 counts in one revolution. The smoothed signal is then feed through a discrete time derivative block which gives the wheel's velocity in rev/sec. A gain then converts the wheel velocity to rad/sec.

The real time controller is limited to sampling at a rate of 20 kHz before memory buffer overflow problems occur. Due to this limit there exists a threshold on the detection of the wheel velocities. Past a wheel velocity of 490 rad/sec aliasing of the wheel occurs. This condition is called the Nyquist criterion, and is equivalent to saying that the sampling frequency must be strictly greater than twice the signal's bandwidth,

$$\frac{20000(\text{samples})}{\text{sec}} \cdot \frac{\text{rev}}{64(\text{period})} \cdot \frac{\text{period}}{4(\text{samples})} \cdot \frac{2\pi(\text{rad})}{\text{rev}} = 490.8750 \left( \frac{\text{rad}}{\text{sec}} \right)$$

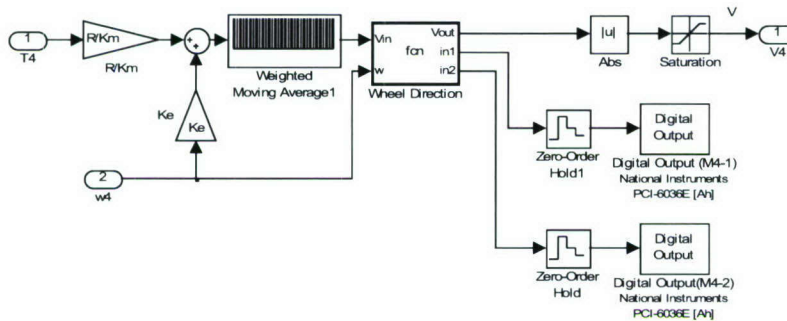
There are 4 samples in each period due to the two encoder channels A and B. If direction was not being accounted for then the wheel velocity detection threshold would be 982 rad/sec; this is a faster speed than the motor can achieve. Since the direction of the motors is important the motor speeds were limited to under 490 rad/sec. Figure 24 shows a motor accelerating to 490 rad/sec, after which aliasing takes place and the wheel speed becomes sporadic.



**Figure 24: Aliasing of the reaction wheel**

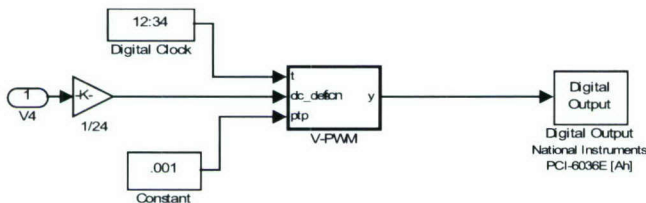
Figure 25 shows the direction and torque subsystem. This subsystem converts the torque and the velocity of the wheel for each motor into the appropriate spin direction and voltage. The voltage to the motor is found by relating the commanded torque to an equivalent commanded voltage. The commanded voltage is found by multiplying the torque by the motor resistance (R) divided by the motor torque constant (Km) and adding it to the voltage created by the spinning armature given by multiplying the motor spin by the back EMF constant (Ke). A running average of the voltage is taken to smooth out the voltage signal. The wheel direction embedded function takes the applied voltage and motor velocity as input and gives the motor direction as output by providing a high or low input to the digital outputs that head to pins IN1 and IN2 of the MOSFET board. The wheel direction embedded code protects

the motor from reversing polarity of the motor terminals when the motor is spinning fast, which would cause a current spike that could destroy the motor coils. The motor is protected at high speeds by first shunting the motor before reversed in direction commanded to the motor. The voltage,  $V_{out}$ , from the wheel direction embedded function is then fed through an absolute value block and a saturation block to make sure the voltage,  $V_4$ , to the motor is both positive and no greater than the max voltage of 24 V.



**Figure 25: Direction and Voltage Torque Subsystem (T2-T4)**

Figure 26 shows the voltage to pulse width modulation (V-PWM) subsystem. The V-PWM subsystem takes the voltage from the direction and voltage torque subsystem ( $V_4$ ) and converts the voltage into an equivalent PWM signal. Inside the V-PWM subsystem is the V-PWM embedded function, which takes the clock running at 20 kHz, a voltage which is limited between 0 (no voltage) and 1 (24 V), and the pulse width period (ptp) at .001, as inputs. The V-PWM gives a pulse width modulated signal at 1 kHz as output to the motors. The PWM signal is fed to the enable pin on the MOSFET board which controls the motor's voltage from 0 to 24 V. The duty cycle of the PWM channel can only be incremented by 5% because 5 is the least common integer value between the 1 kHz PWM frequency and 20 kHz output frequency of the digital output.



**Figure 26: Voltage to Pulse Width Modulation Subsystem (PWM2-PWM4)**

#### IV. Experimental Results of the Cal Poly Spacecraft Attitude Dynamics Simulator

##### A. Simulator Test Setup

The CP/SADS was initially tested with two different maneuvers. First a 90 degree rotation about the z-axis of the CP/SADS was performed. The goal was to show that the sensors and actuator were performing adequately and to demonstrate the initial point capability of the CP/SADS. The second test involved disturbing the CP/SADS with an external torque and commanding the CP/SADS to return to its original orientation. This test demonstrates the ability of the CP/SADS to respond to external torques and allows for the study of the motor saturation characteristics. The CP/SADS was set up with the real time controller as described in the above sections for both tests.

##### B. Test 1: Rotation about a Single Axis

Two tests were conducted on the PRWP using the real time and simulated controllers to show controllability of the platform. The first test required the CP/SADS to maneuver  $-90^\circ$  about the z axis. The initial quaternion for the CP/SADS was set to

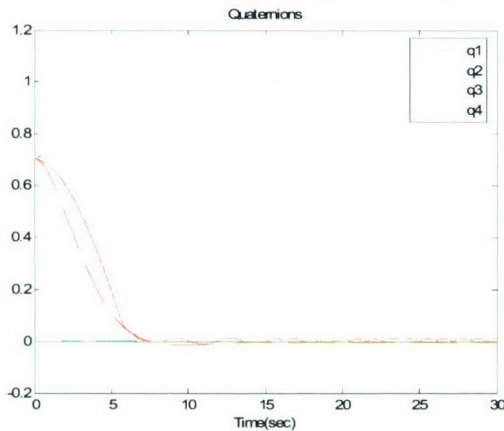
$$\vec{q} = \begin{bmatrix} 0 & 0 & \sqrt{2}/2 \end{bmatrix}^T$$

$$q_4 = \sqrt{2}/2$$

which is a  $90^\circ$  rotation around the Z axis in the inertial frame. The commanded quaternion was set to  $\vec{q}_C = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$  and  $q_{C4} = 1$  which is  $0^\circ$  in the inertial frame. The rotation between the initial and commanded quaternions is a  $-90^\circ$  maneuver in the z body axis.

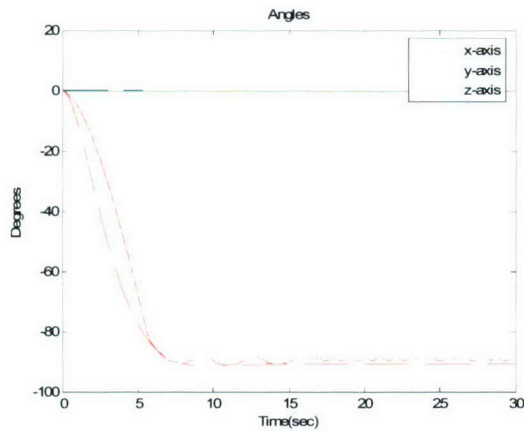
For the simulated controller, K and C gains from Eq. (31) were found that command the platform to its final position quickly and with little overshoot. These same K and C gains were used in the real time controller but altered slightly to match the simulated controller. Differences in the gains of the two controllers are due to two factors. The first is that there are differences between the actual and simulated inertias and the second is that the real time controller is not fully modulated like the simulated controller. The differences between full modulation and finite pulse width modulation are especially pronounced in the small angle steady state maneuvers after the platform has maneuvered itself to its commanded position.

Figure 27 shows the real quaternions of the platform compared to the simulated quaternions (dashed lines) for the first test where the platform maneuvers itself to  $-90^\circ$ . The K gain for the simulated controller was set to 1.1 whereas the real time controller had a K gain of 0.9 making the real time controller slightly more sluggish at the beginning of the maneuver. The C gain for the simulated controller was set to 0.2 whereas the real time controller had a C gain of 0.3 causing the real time controller to dampen out faster. It is also speculated that the current draw from the batteries was saturated during the initial part of the maneuver.



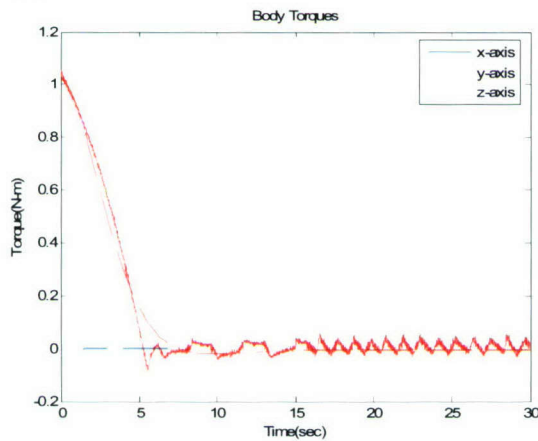
**Figure 27: Quaternions of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 28 shows the corresponding body angles, which are the same as the Euler angles. Since the Z inertial axis is in line with the z body axis the angles in Figure 28 correspond to both the inertial and body axes. The platform approaches  $-90^\circ$  in about 7 seconds.



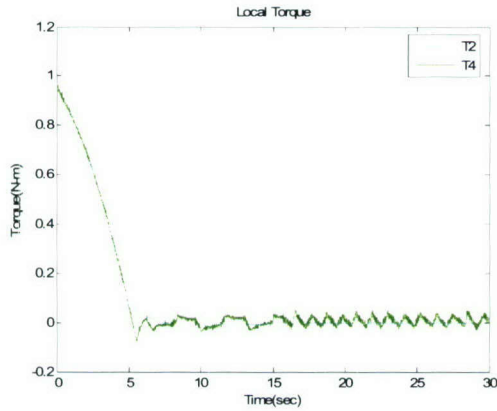
**Figure 28: Euler angles of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 29 shows the body commanded torques from the quaternion control law subsystem. The x and y axes torques are zero because the platform is only maneuvering in the z body axis. Since the real time controller is not fully modulated the real time torques after the platform has reached its commanded position are shifted back and forth meaning the real wheel speed needed to maintain this position is somewhere between the incremented wheel speeds. This phenomenon is known as clocking. The platform compensates for this by switching torques in rapid procession.



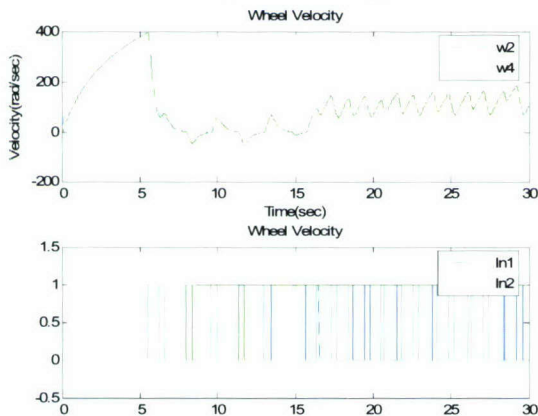
**Figure 29: Body torques of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 30 shows the local wheel torques applied to each wheel from the distributive torque subsystem. Since rotations were only in the z body axis all four local torques of the simulated controller are equal as well as both local torques of the real time controller. The real local torques are twice the simulated local torques. The multiplier of two adjusts the real time controller since it only has half the number of reaction wheels as the simulated controller.



**Figure 30: Local wheel torques of the simulated platform (dashed lines) and actual platform (solid lines)**

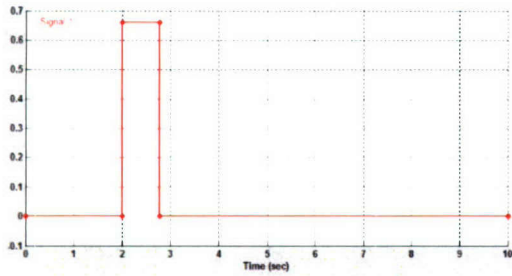
Wheel velocities for the simulated and real platform can be seen in Figure 31. The wheel velocity for the simulated platform starts at zero and after the maneuver is accomplished the ending wheel velocity is again zero. The simulated controller assumes no friction therefore momentum is exchanged directly to the platform. Since there is neither loss in energy nor any external torques the momentum from the platform is transferred back to the wheel to eventually stop the platform. This momentum is equal to the momentum it took to accelerate the platform which causes the ending velocity of the wheel to be zero. Since the real platform does have friction, both in the air bearing and in the reaction wheel (due to the air), the final velocities do not go to zero but are left at around 80 rad/sec.



**Figure 31: Wheel velocities (top) of the simulated platform (dashed lines) and actual platform (solid lines) and direction inputs to the MOSFET board for the actual platform (bottom)**

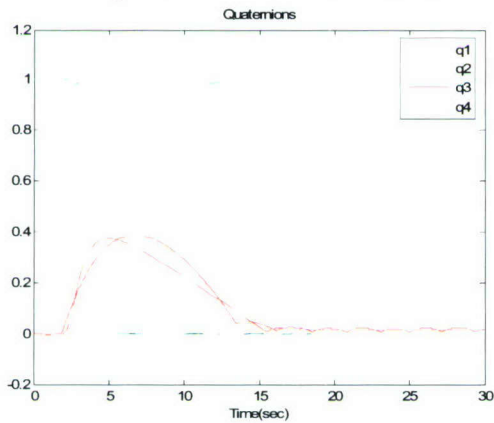
### C. Test2: Disturbance Response

A second test was conducted to show platform stationary maintenance ability with the introduction of a disturbance. Both initial and commanded quaternions were set to  $\vec{q} = [0 \ 0 \ 0]^T$  and  $q_4 = 1$  for this test. In the second test an external moment was applied to the platform to see how the platform would respond. This external moment for the real time controller came in the way of a slight hand nudge in the z body frame whereas the external moment for the simulated controller tried to mimic this hand nudge using the square waveform shown in Figure 32. The real time controller test was completed first and the simulated controller test was designed to imitate the real time controller test by changing the amplitude and period of the square wave.



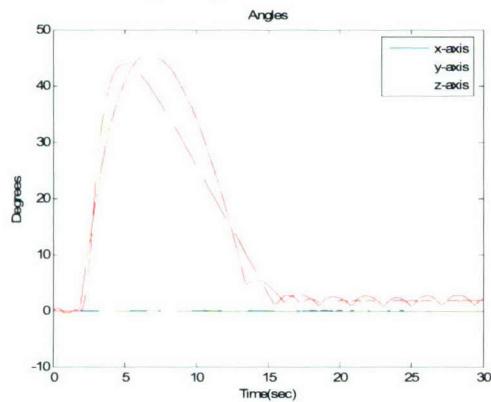
**Figure 32: External moment applied to the simulated platform in the z body axis intended to match moment applied to the actual platform.**

Figure 33 shows the quaternions of the real and simulated controllers when an external moment is applied in the z body axis to the platform 2 seconds into the test.



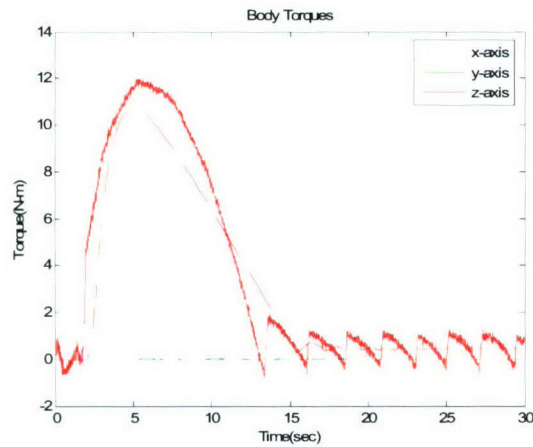
**Figure 33: Quaternions of the simulated platform (dashed lines) and actual platform (solid lines)**

Figure 34 shows the corresponding Euler angles of the real and simulated platforms. The platform is pushed about  $45^\circ$  off its original position before it recovers.

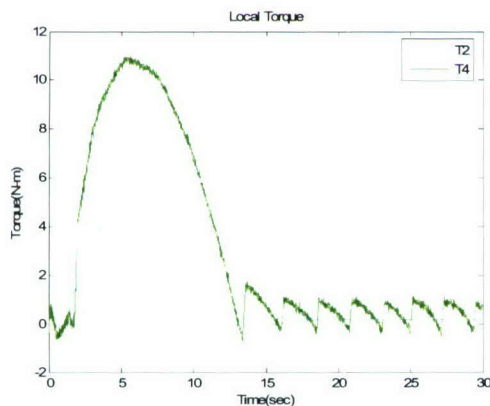


**Figure 34: Euler angles of the simulated platform (dashed lines) and actual platform (solid lines)**

Figures 35 and 36 show the body and local commanded torques of the platform. The real time controller local torques are twice that of the simulated torques because the real time platform only modulates two wheels instead of four on the simulated platform.

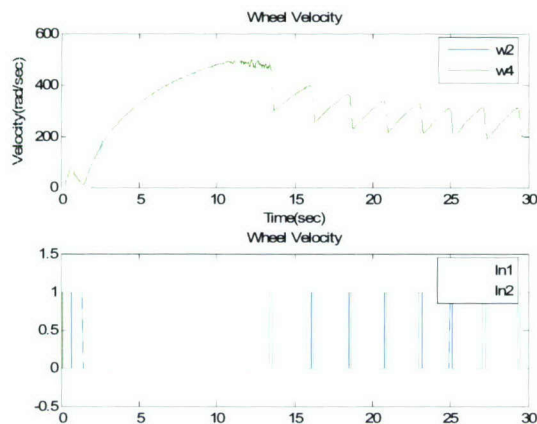


**Figure 35: Body torques of the simulated platform (dashed lines) and actual platform (solid lines)**



**Figure 36: Local wheel torques of the simulated platform (dashed lines) and actual platform (solid lines)**

Wheel velocities for the real and simulated platform are shown in Figure 37. Unlike the first test, the second test of the platform applies an external moment. The wheels increase velocity to combat the positive moment applied. The opposing moment the wheels produce causes the platform to decelerate and eventually accelerate in the opposite direction to get back to its original orientation. Upon arriving at its original position some of the momentum from the platform is transferred back to the wheels but due to the external moment which was applied to the platform the wheels need to keep a constant final velocity. The energy in the spinning wheels at the constant final velocity is proportional to the energy it took to nudge the platform from its initial position. If the platform were given a greater external moment or a series of successive external moments the reaction wheels would reach a limit in which they could no longer produce the opposing torque needed to combat the external torque. When this happens the reaction wheel is said to have saturated. The reaction wheel's saturation speed is the motor's maximum speed which is 710 rad/sec for the Faulhaber motors.



**Figure 37: Wheel velocities (top) of the simulated platform (dashed lines) and actual platform (solid lines) and direction inputs to the MOSFET board for the actual platform (bottom)**

## V. Conclusions, Recommendations and Future Work

The initial development of the CP/SADS has been successful. Both tests described above demonstrate the ability of the CP/SADS to perform as a spacecraft simulator about a single axis. The first test demonstrates the pointing accuracy of the CP/SADS. The pointing accuracy of the CP/SADS needs to be fully characterized, but appears to be around 2-3 degrees.

The second test verifies the well known phenomenon which occurs with momentum exchange devices. Namely, external moments can cause reaction wheel saturation. While the applied external moment was not precise, the concept is surely demonstrated.

The initial development of the CP/SADS has proven to be a very useful learning experience for both students and faculty involved with the project. From designing subsystem hardware to writing real time control algorithms, it was the first time many of these tasks were performed by any of the team members involved. The development of the CP/SADS has provided an invaluable set of institutional knowledge for the students and faculty at Cal Poly.

There are several obvious required improvements before the CP/SADS is truly ready for space system simulation and testing. To complete the development of the CP/SADS, the following items are considered high priority:

- Design and manufacture of the Fine Balancing System
- Integration of the wireless data acquisition and commanding system
- Integration of the Northrop Grumman LN-200 Inertial Measurement Unit

Before the CP/SADS can emulate a fully agile spacecraft, the CP/SADS must be balanced about all three axes. The Fine Balancing System, encompassing the Center of Gravity System Identification algorithm and the Fine Balance Mass Tuning System is critical to collocating the CP/SADS center of gravity and the pivot point of the spherical air bearing. A considerable effort involving many students and faculty is currently underway.

To further enhance ability of the ability of the CP/SADS to emulate the microgravity environment of space, the umbilical cord of wires used to command and control the CP/SADS has to be removed. Hence, the CP/SADS has to move to a wireless data acquisition system. Several hardware components have been identified to design and build such a system. Several students and faculty from the Computer Engineering Program within the Computer Science Department at Cal Poly have been brought on board to work with students from the Aerospace Engineering Department to complete the wireless data subsystem project.

With a balanced simulator which has a wireless data acquisition system, the CP/SADS will be ready for the integration of the Northrop Grumman LN-200 Inertial Measurement Unit. The LN-200 is critical to the CP/SADS since it is capable of independently providing both angular and angular rate measurements. The LN-200 is a much more precise instrument than the current set of MEMS rate gyros used by the CP/SADS. The LN-200 also contains a set of accelerometers which can be used to find angular accelerations of the CP/SADS without taking derivatives of the angular velocity. The angular acceleration is used to identify the center of gravity and inertial characteristics of the CP/SADS.

Once the three items listed above are complete, the CP/SADS will be ready for more projects and tests. The CP/SADS will be capable of simulating various control laws, testing other sensors and actuators, and testing

conceptual designs for future satellite missions. The goal of the CP/SADS platform is to provide a test bed for future satellite attitude control algorithms, sensors and actuators. Our group intends to use the CP/SADS to demonstrate fine pointing control of micro-class satellites for use with optics based systems. We are currently in the conceptual design phase of a small space based optical telescope.

Our group also intends to focus on several areas of fruitful areas of research and engineering design and development including:

- Adaptive Disturbance Rejection of Tonal Reaction Wheel Assemble Disturbances
- IMU, Rate Gyroscope and GPS Sensor Fusion for Attitude Determination and Control
- Linear and Nonlinear System Identification of Reaction Wheel Systems
- Line-of-sight Geolocation for Remote Sensing Space Systems
- Autonomous Control of Satellite Formations
- PolySat Rate Gyro Characterization and Calibration
- Coupled Reaction Wheel and Control Moment Gyroscope Attitude Control

Each of these topics is an exciting area of current research or interest to both the academic and industrial community. By providing an environment for students to learn cutting edge satellite navigation and attitude control techniques, Cal Poly will help our students become technical leaders in industry upon graduation. The CP/SADS and its associated laboratory will also provide academic and industry partners a place to test, develop and integrate future space based navigation and control components.

## VI. Acknowledgements

This work was sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855.

## VII. Bibliography

1. Wie, Bong, *Space Vehicle Dynamics and Control*, Reston, VA, 1998.
2. Christian, John, Turbe, Michael, Kabo, Eric, Manno, Laura, and Johnson, Eric, "Development of a Variable Inertia Reaction Wheel System for Spacecraft Attitude Control," *AIAA Guidance, Navigation and Control Conference*, AIAA-2004-5132, Providence, Rhode Island, August 2004.
3. Rowell, Derek, Wormley, David N., "State Equation Formulation," *System Dynamics*, New Jersey, Prentice Hall, 1997.
4. Healy, Patrick, "Mass Property System Identification of a Spacecraft Simulator," California Polytechnic State University, San Luis Obispo, Master's of Science Thesis, 2006.
5. Martin, Joe, Quinto, Adrian, Nash, Amanda, "Reaction Wheel Platform (Satellite Simulator Structure)," California Polytechnic State University, San Luis Obispo, Senior Project, 2006.
6. Pearce, C, "The Performance and Future Development of a MEMS SIVSG and its Application to the SIIMU," AIAA 2001-4410, 2001.
7. Wiesel, William E., "Rigid-Body Dynamics," *Spaceflight Dynamics*, 2<sup>nd</sup> ed., Irwin McGraw-Hill, Massachusetts, 1995.
8. Huerta, Christopher, Derek, "Development of a Highly Integrated Communication System for use in Low Power Space Applications," California Polytechnic State University, San Luis Obispo, Master's of Science Thesis, 2006.
9. Mehiel, Eric A., and Balas, Mark J. "Nonlinear Direct Model Adaptive Control", *Submitted, Journal of Guidance, Dynamics, and Control*.
10. Space Systems Simulation Laboratory, Department of Aerospace Engineering, Virginia Polytechnic Institute and State University, January 2006. <<http://www.sssl.aoe.vt.edu/dsacss/index.php>>.
11. Tanygin, Sergei, and Trevor Williams. "Mass Property Estimation Using Coasting Maneuver." *Journal of Guidance, Control, and Dynamics*, Jul. - Aug. 1997: 625-632.

# Nonlinear Direct Adaptive Control and Disturbance Rejection for Spacecraft

Simon C. Torres<sup>\*</sup> and Eric A. Mehiel<sup>†</sup>

*California Polytechnic State University, San Luis Obispo, CA, 3407*

**In this paper, a Nonlinear Direct Model Reference Adaptive Control (NDMRAC) is derived. The NDMRAC controller is compared to the Full State Feedback (FSFB) controller. Both of the controllers are applied to a rigid body spacecraft. To compare the controllers, the inertia matrix is suddenly changed in the simulation. Euler equations are used to estimate the evolution of the rigid body angular velocity and quaternions are used to describe the attitude position of the rigid body. The system is augmented or modified to account for the disturbances affecting the system under observation, so the NDMRAC control also implements a Direct Adaptive Disturbance Rejection (DADR) control which partially or fully eliminates the disturbance coming into the simulated system. The error of the system and the power spectrum density of the disturbance are used to analyze the performance of the NDMRAC and DADR controllers.**

## I. Introduction

Since humans have envisioned exploring, there has always been a need to control the spacecraft. Adaptive control is one possible modern control technique since the controller can evolve to the surrounding environment or disturbances coming into the control system. Adaptive control is used to control many dynamic systems that have changing parameter. For instance, sloshing fuel in a spacecraft or flexible appendages can quickly change the inertia matrix of a spacecraft. Thus the applied controller should adapt to or accommodate to sudden unexpected changes. Adaptive control can be used to estimate the control law parameters based on the measured system signals. The Direct Model Reference Adaptive Controller, (DMRAC), attempts to constantly track a given reference model thereby making system performance predictable regardless of the changing system parameters.

Adaptive control has been used for many different applications. Mehiel and Balas<sup>1</sup> apply a DMRAC and DADR control to a deployable telescope to deal with vibration disturbances and fine pointing mission requirements. Barkana and Kaufman<sup>2</sup> were among the first to develop DMRAC control laws for large structural systems with multiple inputs and outputs. Hsu, et al.<sup>3</sup> also derives a MIMO adaptive control algorithm using a high frequency gain in the adaptive algorithm. Barkana<sup>4</sup> also applied adaptive control theory to autopilot design mixed with classical theory (frequency). There are also some applications of adaptive control for spacecraft maneuvers. Junkins and Akella<sup>5</sup> derive adaptive tracking laws to perform a near optimum time maneuvers. Cristi, Burl, and Russo<sup>6</sup> derive adaptive laws to estimate the inertia matrix of the spacecraft, they use quaternions to avoid singularities, and Costic and Dawson<sup>7</sup> further expand on the subject by deriving an adaptive full state feedback controller that requires no angular velocity measurements. They used a filter to generate a velocity related signal from attitude measurements. Their work differs from the work done in this paper since the algorithms derived herein follow the command tracker generator (CGT) approach plus a NDMRAC is developed.

Broussard and O'Brien<sup>8</sup> were among the first to introduce the CGT method, which was further expanded on by Sobel and Kaufman<sup>9</sup>. Later, Su and Sobel<sup>10</sup> derived three different algorithms using the CGT method; one algorithm derived was for feedback, another one for parallel, and the last one for cascade supplementary dynamics.

In order to apply a DMRAC controller, the system should comply with some requirements. One of the requirements is that the plant be Strictly Positive Real (SPR). Ozcelik and Kaufman<sup>11</sup> derived an adaptive controller which satisfies an almost SPR plant. In their work, they add a Feedforward compensator, which alleviates plants which are not SPR and the resulting system is now called almost SPR or ASPR. The derived controller applies to single input single output (SISO) systems. Later on, Ozcelik and Kaufman<sup>12</sup> further expanded the theorem by deriving an algorithm applicable to MIMO systems. Although, the stability analysis is mostly proved through

<sup>\*</sup> Student, Aerospace Engineering Department, AIAA Student Member.

<sup>†</sup> Assistant Professor, Aerospace Engineering Department, AIAA Member.

Lyapunov's stability theory, some people opted for different methods of showing the stability of the system. Monopoli<sup>13</sup> was among the first to introduce a Lyapunov quadratic function to show asymptotic stability using DMRAC controllers. Chuan and Fagan<sup>14</sup> are among the groups which show stability of an MRAC system using Hyperstability instead of the Lyapunov stability analysis. Direct adaptive control algorithms differ from method to method but they all share a common goal which is to track a predetermined model.

No controller is complete without some kind of disturbance rejection control since disturbances are a very important factor in the control design. Disturbances may introduce unwanted behavior in the otherwise well behave system. Johnson<sup>15</sup> derived the theory for disturbance rejection and categorizes the disturbances as noise type disturbances and waveform disturbances. Kelly and Johnson<sup>16</sup> recap the previously derived disturbance rejection theory and developed some extensions of the theorem and apply them to steady state conditions of a system. Fuentes and Balas<sup>17</sup> analyze the conditions which should be satisfied by an MRAC system for it to converge under a set of persistent disturbances. Fuentes, Balas and Erwin<sup>18</sup> further expand on the subject by deriving a discrete time algorithm. In Ref. 19, Mehiel applies the adaptive control and adaptive disturbance rejection control laws derived by Fuentes and Balas to a flexible space structure with the goal of developing a systematic DMRAC design process. Even though the disturbance rejection algorithms may differ, they all have a common goal which is to completely eliminate or to minimize the affect of the disturbances entering into the system.

## II. NDMRAC and DADR Theory

All of the DMRAC theory deals with linear time invariant (LTI) systems, so a Nonlinear Direct Model Reference Adaptive Control (NDMRAC) is derived in this paper expanding on the work done by Fuentes and Balas<sup>17</sup>. The derived NDMRAC controller is applied to a rigid body spacecraft rather than to a flexible space structure. Fig. 1 shows the scheme for the NDMRAC and DADR controllers.

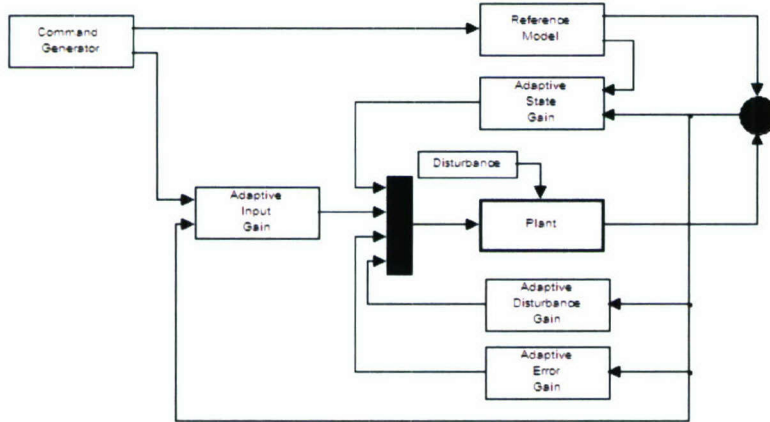


Figure 1 DMRAC and DADR controller block diagram

Consider the  $N^{\text{th}}$  order nonlinear system shown below where  $f(x_p)$  contains the nonlinearities of the system.

$$\dot{x}_p = A_p x_p + B_p u_p + f(x_p) \quad (1)$$

$$y_p = C_p x_p$$

Where  $x_p^0 \equiv x_p(0) \in R^{N_p}$  and if  $\dot{x}_p = f(x_p) + B_p u_p$ ,  $A_p$  is the Jacobian,  $Df(x_p)$ , of  $f(x_p)$  and  $u_p, y_p \in R^M$ , i.e., the system is square. The system in Eq. is said to be output feedback stabilizable if there exists some output feedback gain  $G_e^*$  such that with the control law,  $u_p = G_e^* x_p$ , the operator,  $A_p + B_p G_e^* C_p$  is exponentially stable.

The nonlinear system must track a reference model such as the one given by

$$\begin{aligned} \dot{x}_m &= A_m x_m + B_m u_m + g(x_m) \\ y_m &= C_m x_m \end{aligned} \quad (2)$$

where  $g(x_m)$  is the nonlinear part of the reference model. Also  $x_m^0 \equiv x_m(0) \in R^m$  and if  $\dot{x}_m = g(x_m) + B_m u_m$ ,  $A_m$  is the Jacobian,  $Dg(x_m)$ , of  $g(x_m)$ , and  $N_m \leq N_p$  which means that the reference model can be of lower dimension than that of the plant. Consider the at least marginally stable Command Generator Tracker (CGT).

$$\begin{aligned}\dot{x}_q &= A_q x_q \\ u_m &= C_q x_q\end{aligned}\quad (3)$$

where  $x_q^0 \equiv x_q(0) \in R^{N_q}$  such that all trajectories are bounded. Now, let us introduce some ideal system of the following form

$$\begin{aligned}\dot{x}_* &= A_p x_* + B_p u_* + f(x_*) \\ y_* &= C_p x_*\end{aligned}\quad (4)$$

The ideal trajectories,  $x_*(t)$  are those such that  $y_* = y_m$  for all  $t \geq 0$  which means that the model reference tracks the ideal system hence forcing the plant to track the ideal trajectories. Assuming that the ideal trajectories are a linear combination of the model reference system, a linear transformation can be written as:

$$\begin{bmatrix} x_* \\ u_* \end{bmatrix} = \begin{bmatrix} S_{11}^* & S_{12}^* \\ S_{21}^* & S_{22}^* \end{bmatrix} \begin{bmatrix} x_m \\ u_m \end{bmatrix}\quad (5)$$

Taking the derivative of the linear transformation and substituting Eqs. (2) and (4) results in the following equations which are called the matching conditions.

$$\begin{aligned}A_p S_{11}^* + B_p S_{21}^* &= S_{11}^* A_m \\ (A_p S_{12}^* + B_p S_{22}^*) C_q &= S_{11}^* B_m C_q + S_{12}^* C_q A_q \\ f(x_*) &= S_{11}^* g(x_m) \\ C_p S_{11}^* &= C_m \\ C_p S_{12}^* C_q &= 0\end{aligned}\quad (6)$$

If the matching conditions are satisfied by some  $S_{ij}$  then it is said that the systems shown in Eqs. (2), (3), and (4) are totally consistent systems. With the result above, the NDMRAC theorem can be stated.

*Theorem 1: If  $(A_p + B_p G_e^* C_p, B_p, C_p)$  is SPR, and the systems given in Eqs. (2), (3), and (4) are totally consistent systems, with control law,*

$$u_p = S_{21} x_m + S_{22} u_m + G_e e_y$$

*with the adaptive gain laws,*

$$\begin{aligned}\dot{S}_{21} &= -e_y x_m^T H_1 \\ \dot{S}_{22} &= -e_y u_m^T H_2 \\ \dot{G}_e &= -e_y e_y^T H_3\end{aligned}\quad (7)$$

where  $H_{1,2,3} \in R^{N_m \times N_m, M \times M, M \times M}$  are the positive definite adaptive parameter matrices give asymptotic stability of the system in Eq. (1).

To show asymptotic stability, first define the output error of the system as

$$e_y \equiv y_p - y_m\quad (8)$$

where  $y_p$  and  $y_m$  are defined as in Eqs. (1) and (2) respectively. Now, let the state error of the system be defined as

$$e \equiv x_p - x_*\quad (9)$$

Taking Eqs. (1) and (4), and substituting them into the derivative of the state error results in,

$$\dot{e} = A_p e + B_p \Delta u + \Delta f. \quad (10)$$

Where  $\Delta u = u_p + u_*$  and  $\Delta f = f(x_p) - f(x_*)$ . Solving for  $\Delta u$  and adding the output feedback gain to account for  $y_p \neq y_m$  at  $t = 0$  gives,

$$\dot{e} = (A_p + B_p G_e^* C_p) e + B_p W + \Delta f. \quad (11)$$

Let  $\Delta G = [G - G^*] = [S_{21} \ S_{22} \ G_e] - [S_{21}^* \ S_{22}^* \ G_e^*]$ ,  $z = [x_m \ u_m \ e_y]^T$ , and  $W = \Delta G z$ . Also, let  $A_c = (A_p + B_p G_e^* C_p)$  which results in,

$$\dot{e} = A_c e + B_p W + \Delta f. \quad (12)$$

The stability analysis of the system is done using the Lyapunov stability theory. Assuming that the system  $(A_c, B_p, C_p)$  is SPR, the Lyapunov equation,

$$V_1 = \frac{1}{2} (e^T P_c e) \quad (13)$$

is used to demonstrate the stability of the system.

Taking the derivative of the Lyapunov equation  $V_1$ , substituting Eq. (12), and making use of the Kalman-Yacubovich (Slotine and Weiping<sup>20</sup>) theorem, gives,

$$\dot{V}_1 = -\frac{1}{2} e^T Q_c e + e_y^T W + e^T P_c \Delta f. \quad (14)$$

Now consider the second Lyapunov of equation of the form

$$V_2 = \frac{1}{2} \text{tr}(\Delta G H^{-1} \Delta G) \quad (15)$$

where  $H$  is a diagonal positive definite matrix. Taking the derivative of the second Lyapunov equation and letting  $\Delta \dot{G} = -e_y^T z^T H$  results in

$$\dot{V}_2 = -e_y^T W \quad (16)$$

Now let  $V = V_1 + V_2$  which gives

$$\dot{V} = -\frac{1}{2} e^T Q_c e + e_y^T W + e^T P_c \Delta f - e_y^T W \quad (17)$$

resulting in

$$\dot{V} = -\frac{1}{2} e^T Q_c e + e^T P_c \Delta f \quad (18)$$

Taking the absolute value of the nonlinear part gives,

$$\dot{V} \leq -\frac{1}{2} e^T Q_c e + |e^T P_c \Delta f|.$$

Using the inequality,

$$|e^T P_c \Delta f| \leq \|e^T\| \|P_c\| \|\Delta f\|$$

results in

$$\dot{V} \leq -\frac{1}{2} e^T Q_c e + \|e^T\| \|P_c\| \|\Delta f\|.$$

If the nonlinearity  $f(x)$  is Lipschitz in some ball,  $\beta(0, r)$  then  $\|f(x_p) - f(x_*)\| \leq L \|x_p - x_*\|$  (Slotine and Weiping<sup>20</sup>) or  $\|\Delta f\| \leq L \|e\|$ . Substituting this result into equation (18) gives,

$$\dot{V} \leq -\frac{1}{2} e^T Q_c e + \|e^T\| \cdot \|P_c\| \cdot L \|e\| \quad (19)$$

Also,  $e^T Q_c e \geq \underline{\sigma}(Q_c) \|e\|^2$  where  $\underline{\sigma}(Q_c)$  is the smallest singular value of  $Q_c$ . Changing sign reverses the inequality giving  $-e^T Q_c e \leq -\underline{\sigma}(Q_c) \|e\|^2$ . Finally, Eq. (19) can be cast as,

$$\dot{V} \leq -\frac{1}{2} \underline{\sigma}(Q_c) \|e\|^2 + \|P_c\| \cdot \|e\|^2 L \quad (20)$$

Let  $\|P_c\| = \bar{\sigma}(P_c)$ , which is the greatest singular value of  $P_c$  resulting in

$$\dot{V} \leq -\|e\|^2 (\underline{\sigma}(Q_c) - 2\bar{\sigma}(P_c)L) \quad (21)$$

If the Lipschitz constant is bounded by,

$$L < \frac{\underline{\sigma}(Q_c)}{2\bar{\sigma}(P_c)} \quad (22)$$

then the derivative of the Lyapunov function,  $\dot{V}$ , is negative semi definite. By Lyapunov theory (Slotine and Weiping<sup>20</sup>), all trajectories are bounded. However, we can not state that the system has asymptotic stability because the function is only a function of the state error. Taking the derivative of Eq. (18) gives

$$\ddot{V} = -e^T Q_c \dot{e} + \dot{e}^T P_c \Delta f + e^T P_c \Delta \dot{f} \quad (23)$$

The state error,  $e$ , is bounded and the matrices  $P_c$  and  $Q_c$  are constant. Recalling  $\dot{e} = A_c e + B_p W + \Delta f$ , and  $W = \Delta G z$ , we remember that  $\Delta G$  is bounded, and  $z$  is a vector formed by the input vector,  $u_m$ , assumed to be bounded, the states of the model reference,  $x_m$ , which are also bounded since the reference model is stable, and the output error  $e_y$ . Remembering that  $e_y = C_p e$ , notice that the output error is also bounded since the output error depends on the state error which is bounded. Based on all of this, it is established that  $z$  is bounded. Assuming that the nonlinearity is Lipschitz,  $\|\Delta f\| \leq L \|\Delta x\|$  where  $\Delta x = x_p - x_* = e$  results in  $L \|\Delta x\| = L \|e\| < \frac{\underline{\sigma}(Q_c)}{2\bar{\sigma}(P_c)} \|e\|$ .

And since  $e$  is bounded  $\Rightarrow \exists M > 0 \ni \|e\| < M$ , also  $\frac{\underline{\sigma}(Q_c)}{2\bar{\sigma}(P_c)}$  is a constant greater than zero which implies

$\|\Delta f\| < \frac{\underline{\sigma}(Q_c)}{2\bar{\sigma}(P_c)} M$  therefore  $\Delta f$  is bounded. Thus far, every term of the Eq. (23) is bounded except for the

derivative of the nonlinear term,  $(\Delta \dot{f})$ . For the analysis here, the nonlinear parts of the dynamic equations are composed of Euler's equations of rotational motion and the kinematic quaternion equations from Eqs. (29) and (30). Looking at the derivatives of the mentioned equations, it is seen they are composed by the trajectories of the system, which are bounded; hence the derivative itself is bounded. Therefore the  $\Delta \dot{f}$  term is bounded. Since every term of Eq. (23) is shown to be bounded, by Barbalat's lemma,  $\dot{V}$  is continuous therefore, the  $\lim_{t \rightarrow \infty} \dot{V} = 0$ . Thus  $\lim_{t \rightarrow \infty} e = 0$

since  $Q_c$  is positive definite. Therefore the output tracking error goes to zero asymptotically while everything else remains bounded. This completes the proof for the Nonlinear Direct Model Reference Adaptive Control (NDMRAC) theory.

Now, augment the system given in Eq. (1) with a disturbance, so the new plant is given by

$$\begin{aligned} \dot{x}_p &= A_p x_p + B_p u_p + f(x_p) + \Gamma_d u_d \\ y_p &= C_p x_p \end{aligned} \quad (24)$$

where  $\Gamma_d$  is the disturbance input matrix, and  $u_d \equiv \sum_{i=1}^l \alpha_i \xi_i \phi_i(t)$ , where  $\xi_i$  are standard basis vectors in  $\mathbb{R}^l$  and  $\phi_i(t)$  are known, bounded functions. Assume exact disturbance cancellation, i.e. that  $\text{span}(\Gamma_d) \subseteq \text{span}(B_p)$ . If  $(A_p + B_p G_e^* C_p, B_p, C_p)$  is SPR, and the systems given in equations (2), (3), and (4) are totally consistent systems, then the adaptive gain laws become

$$\begin{aligned}\dot{S}_{21} &= -e_y x_m^T H_1 \\ \dot{S}_{22} &= -e_y u_m^T H_2 \\ \dot{G}_e &= -e_y e_y^T H_3 \\ \dot{H}_d &= -e_y \phi_d^T H_4\end{aligned}\tag{25}$$

Where  $H_i$  are positive definite matrices, and are called the adaptive parameters along with the control law, which is now given as  $u_p = S_{21}x_m + S_{22}u_m + G_e e_y + H_d \phi_d$  produces asymptotic tracking with uniformly bounded adaptive gains.

Since the span of  $\Gamma_d$  is contained in the span of  $B_p$ , a transformation  $T$  exists such that  $\Gamma_d = B_p T$ . Letting

$E \equiv -[\alpha_1 \xi_1 \quad \alpha_2 \xi_2 \quad \cdots \quad \alpha_l \xi_l]$  and  $H_d^* \equiv T \cdot E$  the plant equation (24) can be rewritten as

$$\begin{aligned}\dot{x}_p &= A_p x_p + B_p (u_p - H_d^* \phi_d) + f(x_p) \\ y_p &= C_p x_p\end{aligned}\tag{26}$$

The output error and the state error are defined as in Eqs. (8) and (9) respectively. Taking the derivative of the state error gives

$$\dot{e} = A_p e + B_p (\Delta u - H_d^* \phi_d) + \Delta f\tag{27}$$

Augmenting  $\Delta G$  to include the adaptive disturbance gain term gives,

$$\Delta G = [S_{21} \quad S_{22} \quad G_e \quad H_d] - [S_{21}^* \quad S_{22}^* \quad G_e^* \quad H_d^*]$$

and  $z$  to include the bounded basis functions,

$$z = [x_m \quad u_m \quad e_y \quad \phi_d]^T$$

equation (27) can be rewritten as

$$\dot{e} = A_c e + B_p W + \Delta f\tag{28}$$

recall that  $W = \Delta G z$ , and the adaptive gain laws can be expressed as  $\Delta \dot{G} = -e_y z^T H$ . Since the derivative of the state error in (28) has the same form as in equation (12), the Lyapunov stability analysis proceeds as before, and all of the theorems used before still apply since the disturbance function,  $\phi_i(t)$ , is bounded. This concludes the proof for the Direct Adaptive Disturbance Rejection (DADR) Theory.

### III. Dynamic System and Simulation Results

The Euler and Kinematic equations form the plant for the simulated system. The Euler equations describe the evolution of the rigid body's angular velocities as time progresses. Quaternions are used as the Kinematic equations. Using quaternions allows a faster on-line calculation, and avoids any singularities that may arise when using Euler angles. The Euler equations which are also called the dynamic equations are given as

$$\vec{M} = \dot{\vec{h}}_l = \dot{\vec{h}} + \vec{\omega} \times \vec{h}\tag{29}$$

Equation (29) is in vector form where  $\vec{M}$  is the external moments applied to the rigid body,  $\vec{h}$  is the angular momentum of the rigid body, and  $\vec{\omega}$  is the angular velocity of the rigid body. The angular momentum  $\vec{h}$  is

defined as  $\vec{h} = J\vec{\omega}$  where J is the Inertia matrix of the rigid body. In this work, the principal moments of inertia are used for all the calculations done during the simulation; thus the inertia matrix is a diagonal matrix. The quaternion equations describe the attitude of the rigid body at all times; this is important to know. Otherwise for instance, a plane can crash or a communications satellite can face its antenna to outer space in both cases ending in failure. As derived by Wie<sup>21,22</sup>, the quaternion equations are

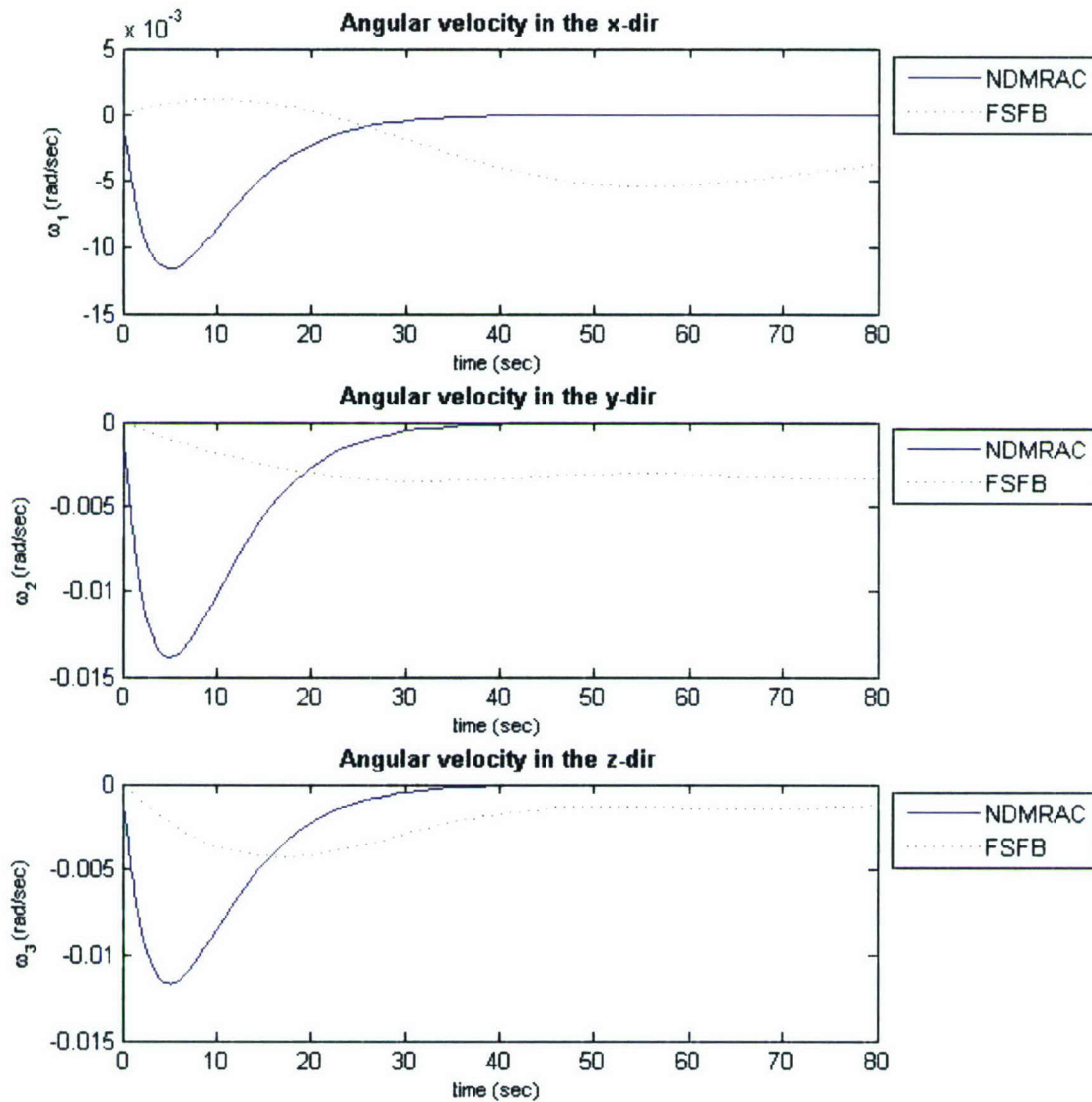
$$\begin{aligned}\dot{\vec{q}} &= \frac{1}{2}(\vec{q}_4\vec{\omega} - \vec{\omega} \times \vec{q}) \\ \dot{q}_4 &= -\frac{1}{2}\vec{\omega}^T \vec{q}\end{aligned}\tag{30}$$

NDMRAC, DADR, and FSFB controllers are applied to the nonlinear system previously discussed. FSFB has been extensively researched by Wie<sup>21,22</sup> and others; the stability for the system using FSFB control has been proven using Lyapunov's stability theory.

$$\vec{M} = -K\vec{q}_e - C\vec{\omega}\tag{31}$$

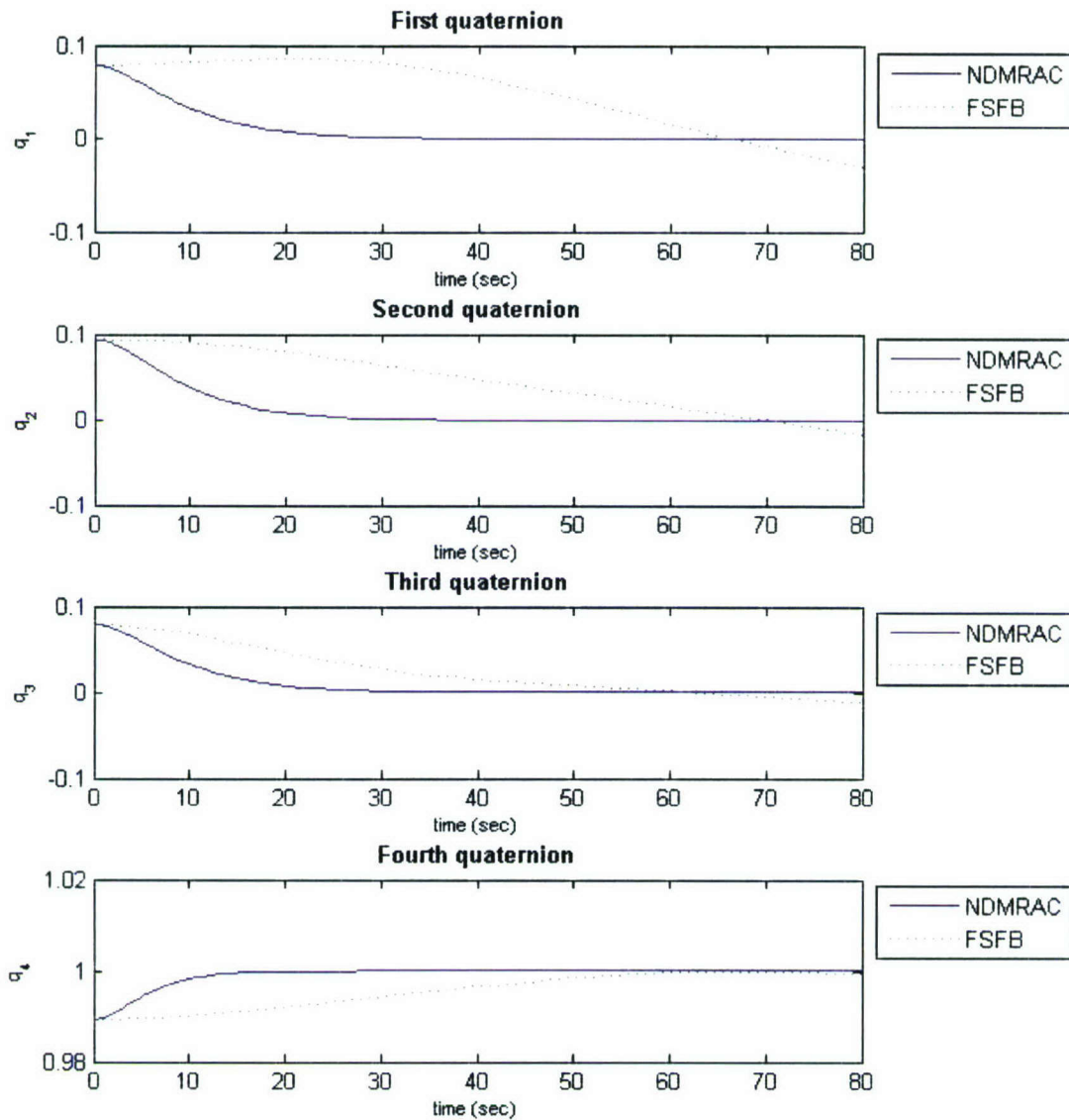
The control law used for the FSFB is represented in Eq. (31) where  $\vec{q}_e$  is the vector part of the quaternion error; this quaternion vector comes from the multiplication of the rigid body quaternion and the commanded quaternion, a derivation of the quaternion error was done by Wie<sup>21</sup> and Sidi<sup>23</sup>, and Kuipers<sup>24</sup> derives the general quaternion control law.

The task of the rigid body is to track the origin from a given quaternion. To compare the NDMRAC controller to the FSFB controller, the inertia matrix of the system changes after 20 seconds into the simulation. Figure 2 shows the angular velocities for both controllers. The angular velocities for the NDMRAC controller start at zero and then go back to zero just as expected. The FSFB controller struggles and cannot make the angular velocities settle down to zero within the simulation time.



**Figure 2 Angular velocities of the spacecraft**

The quaternions for the NDMRAC and the FSFB controller are shown in Figure 3. The quaternions obtained from the NDMRAC controller perfectly track the origin while the quaternions from the FSFB controller do not. This result is somewhat expected since the angular velocities are used to calculate the quaternions. Thus with these results, it is demonstrated that the NDMRAC controller works better than the FSFB controller for spacecraft systems when the inertial matrix changes suddenly due to some shift in mass. The shift in mass could represent a solar panel or sensor repositioning or some failure mode. The NDMRAC controller adapts to the unexpected parameter changes making it possible to control the rigid body. On the other hand, the FSFB controller lacks such an ability making it difficult for the rigid body to perform the prescribed maneuver without some sort of gain scheduling or other technique.



**Figure 3 Quaternions of the spacecraft**

Once the NDMRAC controller was demonstrated to work better than the FSFB controller, a disturbance of the form  $\sin(10\pi t) + \cos(7\pi t)$  is added to the system. Figure 4 shows the error of the system when the disturbance adaptive gain is zero. The disturbance is not rejected from the system. Figure 5 on the other hand, shows the disturbance being rejected from the system. The disturbance adaptive gain is no longer zero therefore the system is able to accommodate the disturbance entering the system. Rejecting the disturbance allows the system to converge to zero as expected from the stability analysis.

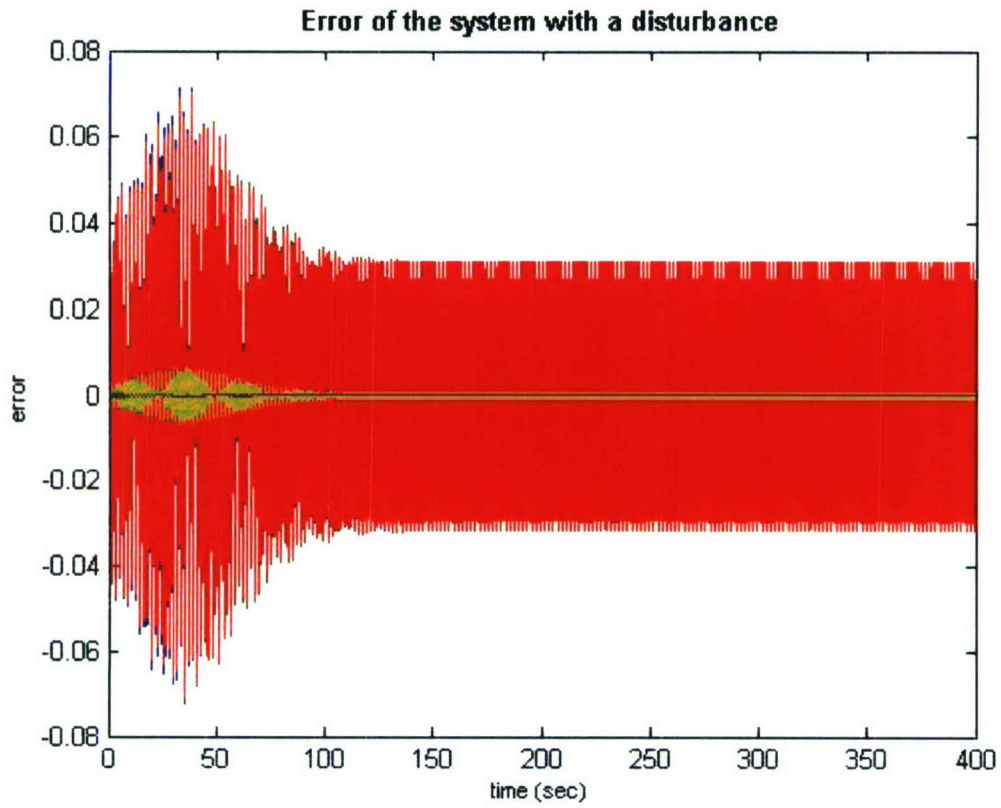
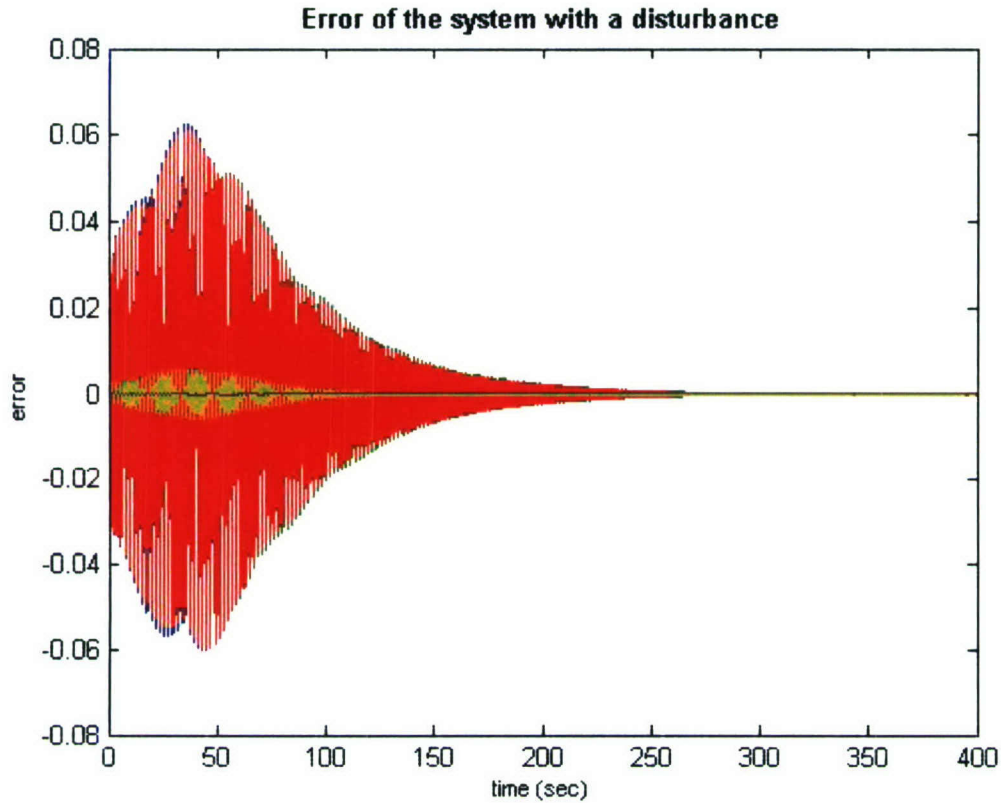
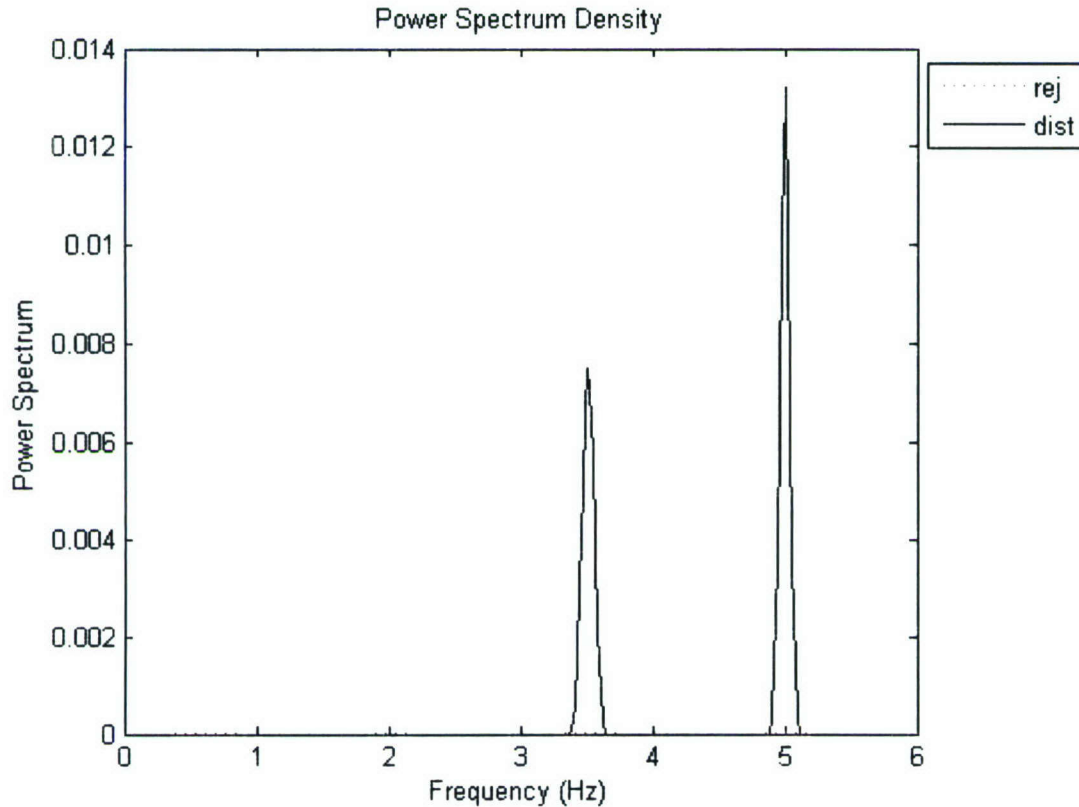


Figure 4 Error of the system



**Figure 5 Error of the system**

Another way to see what is going on with the disturbance is to look at the power spectrum density (PSD) of the disturbance. The PSD is the amount of power under certain frequencies. Figure 6 shows the PSD for the disturbance, and the PSD for the rejection of the disturbance. The PSD graph of the rejected disturbance is below the PSD graph of the disturbance. The result makes sense since the disturbance has more power than when it is rejected. These results demonstrate that the NDMRAC and DADR controllers work well as it was previously predicted by the stability analysis.



**Figure 6 Power spectrum density of the disturbance**

#### IV. Conclusion

The goal of this work was to derive a Nonlinear Direct Model Reference Adaptive Controller (NDMRAC) to deal with nonlinear plants because the DMRAC controller controls linear plants. The derived NDMRAC controller was compared to the FSFB controller. Once the NDMRAC was demonstrated to work better than the FSFB controller, a direct adaptive disturbance rejection (DADR) controller was implemented along with the NDMRAC controller. With this, the controller is more complete because it not only controls the system, but can also reject or accommodate disturbances coming into the system.

To test the ability of the NDMRAC controller and the FSFB controller to adapt, the inertia matrix was changed. The graphs for the simulation with the inertia matrix changed, showed that the NDMRAC controller performs well achieving the prescribed maneuver. The same cannot be said about the FSFB controller. With the parameters changed, the FSFB controller losses it, and cannot track the origin as well as the NDMRAC controller did. From this, it is deduced that an adaptive controller works better than a fixed controller does under unexpected parameter changes. So the adaptive controller manages to adapt any parameter changes making it possible for the system to still be controllable, and unfortunately the FSFB controller cannot since it is a fixed controller lacks the ability to adapt.

A disturbance was fed into the system to test the DADR controller along with the NDMRAC controller. The output error of the system was analyzed for two cases, with no disturbance rejection and disturbance rejection. Further analysis was done using power spectrum density data for both cases. The power spectrum density decays as the disturbance is rejected. The controller completely rejects the disturbance and asymptotically converges to zero.

#### References

<sup>1</sup>Mehiel, A. Eric, and Mark J. Balas. "Adaptive Control for a Deployable Optical Telescope." *AIAA Guidance, Navigation, and Control Conference and Exhibit* (2004): 1-14.

- <sup>2</sup>Barkana, I., H. Kaufman. "Some Applications of Direct Adaptive Control to Large Structural Systems." *Journal of Guidance* Volume 7 (1984): 717-724.
- <sup>3</sup>Hsu, Liu, et al. "Lyapunov-based adaptive control of MIMO systems." *Proceedings of the American Control Conference* (2001): 4808-4813.
- <sup>4</sup>Barkana, Itzhak. "Classical and Simple Adaptive Control for Nonminimum Phase Autopilot Design." *Journal of Guidance, Control and Dynamics* Volume 28 (2005): 631-638.
- <sup>5</sup>Junkins, L. John, Maruthi R. Akella, and Rush D. Robinett. "Nonlinear Adaptive Control of Spacecraft Maneuvers." *Journal of Guidance, Control and Dynamics* Volume 20 (1997): 1104-1110.
- <sup>6</sup>Cristi, Roberto, Jeffrey Burl, and Nick Russo. "Adaptive Quaternion Feedback Regulation for Eigenaxis Rotations." *Journal of Guidance, Control and Dynamics* Volume 17 (1994): 1287-1291.
- <sup>7</sup>Costic, T. B., et al. "Quaternion Based Adaptive Attitude Tracking Controller Without Velocity Measurements." *Journal of Guidance, Control and Dynamics* Volume 24 (2001): 1214-1222.
- <sup>8</sup>Broussard, R. John, and Mike J. O'Brien. "Feedforward Control to Track the Output of a Forced Model." *IEEE Transactions on Automatic Control* Volume AC-25 (1980): 851-853.
- <sup>9</sup>Sobel, M. Kenneth, and Howard Kaufman. "Direct Model Reference Adaptive Control for a class of MIMO Systems." *Control and Dynamic Systems* Volume 24 (1986): 246-311.
- <sup>10</sup>Su, Wei, and Kenneth M. Sobel. "A unified Theory for the CGT Approach to Adaptive Control." *Proceedings of the 30<sup>th</sup> Conference on Decision and Control* (1991): 2915-2916.
- <sup>11</sup>Ozcelik, Selahattin, and Howard Kaufman. "Robust Direct Model Reference Adaptive Controllers." *Proceedings of the 34<sup>th</sup> Conference on Decision and Control* (1995): 3955-3960.
- <sup>12</sup>Ozcelik, Selahattin, and Howard Kaufman. "Design of MIMO Robust Direct Model Reference Adaptive Controllers." *Proceedings of the 36<sup>th</sup> Conference on Decision and Control* (1997): 1890-1895.
- <sup>13</sup>Monopoli, V. Richard. "Lyapunov's Method for Adaptive Control System Design." *IEEE Transactions on Automatic Control* (1967): 334-335.
- <sup>14</sup>Chuan, Zhao, and John E. Fagan. "Model Reference Adaptive Control System Hyperstability Analysis." *IEEE* (1993): 307-309.
- <sup>15</sup>Johnson, C. D. "Theory of Disturbance Accommodating Controllers." *Control and Dynamic Systems* Volume 12 (1976): 387-487.
- <sup>16</sup>Kelly, W. C., and C. D. Johnson. "Theory of Disturbance Utilizing Control; Some Recent Developments." *IEEE* (1981): 614-620.
- <sup>17</sup>Fuentes, J. Robert, and Mark J. Balas. "Direct Adaptive Rejection of Persistent Disturbances." *Journal of Mathematical Analysis and Applications* (2000): 28-39.
- <sup>18</sup>Fuentes, J. Robert, Mark J. Balas, and R. Scott Erwin. "Direct Adaptive Disturbance Rejection and Control for a Deployable Space Telescope, Theory and Application." *Proceedings of the American Control Conference* (2001): 3980-3985.
- <sup>19</sup>Mehiel, E. A. *On Direct Model Reference Adaptive Controller Design for Flexible Space Structures*, Ph. D. Dissertation, Aerospace Engineering Sciences, University of Colorado, Boulder, Colorado, 2003.
- <sup>20</sup>Slotine, E. Jean-Jacques, and Weiping Li. *Applied Nonlinear Control*. Englewood Cliffs: Prentice Hall, 1991.
- <sup>21</sup>Wie, Bong. *Space Vehicle Dynamics and Control*. Reston: AIAA, Inc., 1998.
- <sup>22</sup>Wie, Bong, and Peter M. Barba. "Quaternion Feedback for Spacecraft LargeAngle Maneuvers." *Journal of Guidance, Control and Dynamics* Volume 8 (1985): 360-365.
- <sup>23</sup>Sidi, Marcel J. *Spacecraft Dynamics and Control: A Practical Engineering Approach*. Cambridge: Cambridge UP, 2000.
- <sup>24</sup>Kuipers, B. Jack., *Quaternions and Rotation Sequences*. Princeton: Princeton University Press, 1999.

**Assessment of contaminant transport into groundwater from selected landfill  
sites in the state of California**

Project Investigator:

Saad Merayyan  
Civil & Environmental Engineering  
California Polytechnic State University  
San Luis Obispo, California

## Table of Contents

1	Introduction.....	407
1.1	History of Landfill Construction .....	407
1.2	Motivation for this Research .....	408
1.3	Governing Equation.....	409
2	Background Information.....	409
2.1	Soil Theory .....	409
2.1.1	Flow through a Soil Element.....	410
2.1.2	Introduction to the Soil Water Characteristic Curve .....	413
2.1.3	SWCC Functions.....	415
2.1.4	Permeability Functions.....	416
3	Literature Review.....	418
3.1	Properties of Clay that may be Influenced by Leachate .....	418
3.2	Recent Investigations.....	419
4	Clay and Leachate Descriptions.....	420
4.1	Leachate.....	420
4.2	Clay.....	422
4.2.1	Geotechnical Investigations .....	422
5	Introduction to Numerical Model and Pressure Plate Extractor .....	425
5.1	Pressure Plate Extractor.....	425
5.2	Numerical Analysis Description.....	428
5.3	Limitations.....	429
5.3.1	Soil Water Characteristic and Permeability Inputs to CHEMFLO .....	430
6	Methodology .....	431
6.1	Laboratory Testing .....	431
6.1.1	Sample Preparation .....	431
6.1.2	Laboratory Testing.....	433
6.2	Parameter Optimization Methods.....	437
6.2.1	Solver .....	439
6.2.2	Parameter Optimization Results.....	441
6.3	CHEMFLO – 2000 Methods .....	442
6.4	Inputs to CHEMFLO .....	444
6.4.1	Initial Conditions.....	448
6.4.2	Initial Boundary conditions.....	449
6.4.3	Final Boundary Conditions .....	450
6.4.4	Mesh Size and Convergence .....	450
7	Results.....	452
7.1	Result 1 – Statistical Significance .....	452
7.2	Result 2 – Soil Water Characteristic Curve.....	453
7.3	Result 3 – Quantitative Comparison of Effects of Leachate .....	458
7.3.1	Modeling Scenario #1: Fixed Conductivity .....	458
7.3.2	Modeling Scenario #2: Independent Conductivity.....	462
7.3.3	Modeling Scenario #3: Normalized $\Theta_s$ , $\Theta_r$ , and Dry Densities .....	466
8	Discussions .....	467
8.1	Affects on the SWCC .....	467
8.2	Modeling Scenario #1.....	468
8.3	Modeling Scenario #2.....	470
8.4	Modeling Scenario # 3.....	471
9	Summary and Conclusions.....	474
9.1	Engineering Implications.....	475

9.2	Future Research .....	476
10	References.....	478

## List of Tables

Table 1: Waste Constituents of Landfill where Leachate was Sampled .....	421
Table 2: Possible Leachate Constituents.....	421
Table 3: Possible Leachate VOC Constituents .....	421
Table 4: Previous Geotechnical Investigation .....	422
Table 5: Standard Compaction Results.....	423
Table 6: Atterberg Limits.....	424
Table 7: Specific Gravity .....	424
Table 8: USCS Classification and Grain Size Distribution .....	425
Table 9: CHEMFLO Conductivity and SWCC Functions.....	430
Table 10: Input Data to Parameter Optimization .....	437
Table 11: Excel Solver Configuration .....	440
Table 12: Parameter Optimization Results .....	441
Table 13: Soil System Input to CHEMFLO for Van Genuchten SWCC.....	445
Table 14: Soil System Input to CHEMFLO for Brooks Corey SWCC .....	447
Table 15: Results of Mesh/Convergence .....	451
Table 16: Experimental Data Results.....	452
Table 17: T-Test Results.....	453
Table 18: Leachate's Affect on BC and VG Parameters .....	468
Table 19: Van Genuchten Difference in Initial Properties.....	468
Table 20: Brooks Corey Difference in Initial Properties .....	469
Table 21: Van Genuchten Difference in Initial Properties.....	470
Table 22: Brooks Corey Difference in Initial Properties .....	470
Table 23: Initial Hydraulic Conductivity and Matric Suction .....	472
Table 24: Normalized $\Theta_s$ , $\Theta_r$ , and Dry Densities.....	472

## List of Figures

Figure 1: Common Soil-Water Characteristic Curve (from Merayyan et. al “Soil Type...”)	414
Figure 2: Standard Proctor Compaction Test Results	424
Figure 3: Pressure Plate Extractor and Pressure Regulator	426
Figure 4: Diagram of the Pressure Plate Extractor	427
Figure 5: Diagram of the PPE Pressure Regulation System	428
Figure 6: Van Genuchten SWCC for Water	454
Figure 7: Van Genuchten SWCC for Leachate	454
Figure 8: Brooks Corey SWCC for Water	455
Figure 9: Brooks Corey SWCC for Leachate	456
Figure 10: Comparison of Brooks Corey and Van Genuchten SWCC for Leachate and Water	457
Figure 11: Comparison of Van Genuchten Predicted Matric Potentials using Fixed Permeability Functions	458
Figure 12: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Fixed Permeability Functions	459
Figure 13: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Fixed Permeability Functions (Before Saturation)	459
Figure 14: Comparison of Van Genuchten Predicted Cumulative Flux using Fixed Permeability Functions	460
Figure 15: Comparison of Brooks Corey Predicted Matric Potential using Fixed Permeability Functions	460
Figure 16: Comparison of Brooks Corey Predicted Hydraulic Conductivity using Fixed Permeability Functions	461
Figure 17: Comparison of Brooks Corey Predicted Cumulative Flux using Fixed Permeability Functions	461
Figure 18: Comparison of Van Genuchten Predicted Matric Potential using Independent Permeability Functions	462
Figure 19: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Independent Permeability Functions	462
Figure 20: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Independent Permeability Functions (Before Saturation)	463
Figure 21: Comparison of Van Genuchten Predicted Cumulative Flux using Independent Permeability Functions	463
Figure 22: Comparison of Brooks Corey Predicted Matric Potential using Independent Permeability Functions	464
Figure 23: Comparison of Brooks Corey Predicted Hydraulic Conductivity using Independent Permeability Functions	464

Figure 24: Comparison of Brooks Corey Predicted Hydraulic Conductivity using Independent Permeability Functions (Before Saturation).....	465
Figure 25: Comparison of Brooks Corey Predicted Cumulative Flux using Independent Permeability Functions.....	465
Figure 26: Comparison of Leachate and water with Normalized $\Theta_s$ , $\Theta_r$ , and dry densities (also showing non-normalized parameters) .....	466
Figure 27: Comparison of Leachate and water with Normalized $\Theta_s$ , $\Theta_r$ , and dry densities (also showing non-normalized parameters) .....	466
Figure 28: Percent Difference between Leachate and Water Volumetric Water Contents.....	467
Figure 29: Hydraulic Conductivity vs. Matric Suction.....	471
Figure 30: Van Genuchten Normalized Volumetric Water Content for Leachate and Water .....	473
Figure 31: Brooks Corey Normalized Volumetric Water Content for Leachate and Water.....	474

## **1 Introduction**

### **1.1 History of Landfill Construction**

Historically solid waste landfills have included natural attenuation and containment landfills. Natural attenuation landfills focused on non-hazardous waste from municipal sources and typically leached into groundwater due to poorly constructed or no liners (Bagachi 1). Containment Landfills were developed in the latter half of the twentieth century to accept wastes from industrial sources in addition to municipal wastes. These landfills used low permeability liners to restrict seepage of leachate to groundwater. As water percolates through the solid wastes, biological materials and chemical constituents are leached into solution as the waste undergoes simultaneous and interrelated biological, chemical, and physical changes. This solution is referred to as leachate (Bagachi 2).

Current practice of sanitary landfill engineering came into play in 1984 with an amendment to the 1976 Resource Conservation and Recovery Act that gave the EPA jurisdiction over landfills (Hickman and Richard 2). The EPA responded with the prescriptive Subtitle D that is still referred to today for design specifications. In California landfills are classified into three categories:

- Class I. Hazardous Waste
- Class II. Designated Waste
- Class III. Municipal Waste

Liners are materials which line the bottom and below grade-sides of a landfill and act as a final barrier to the underlying hydrogeology. Their primary function is to control leachate. Liners collect the landfill infiltration and divert it into leachate collection systems so standing water forms above the liner, forming a hydraulic head. Liners are the final defense for underlying aquifers. Because liners are subject to leachate infiltration, they are the focus of this research.

Covers and liners are multilayered systems with the focus around a low permeability layer. They are made up of three general materials: Geomembrane, geosynthetic clay liner, and compacted clay liners. Often a composite liner is used which utilizes a

combination of these materials. Still today clay continues to be an important material in landfill liner design. When clay is used as a hydraulic barrier, field compaction usually occurs at a moisture content wet of optimum. This tends to yield a more oriented or dispersed structure which exhibits a lower hydraulic conductivity. Moisture contents ranging from 2% to 4% above optimum are common (Bagachi 241).

Both Federal and California law mandate cover and bottom liner systems for Class I, II, and III. Final Covers are closure and post closure measures which mostly serve to minimize infiltration into the landfill, mitigate erosion, and control gas emissions. Interim covers are spread over the active landfill face in order to control vectors, fire, water infiltration, erosion, and to prevent unsightliness (Dunn and Singh 10). Although cover liners are not naturally subjected to leachate seepage, facilities that use leachate recirculation systems will pump leachate from above the bottom liner and redirect it to above the waste layer or to a treatment system where it is removed from the landfill.

Minimum requirements for cover and liner systems are outlined in Titles 40, Part 258 of the Code of Federal Regulations and in Title 27, Chapter 3 of California Code of Regulations. Liners are required to meet minimum standards outlined in Subtitle D of these documents.

## **1.2 Motivation for this Research**

Federal and state regulations prescribe design requirements for low permeability layers based on intrinsic hydraulic conductivity using water as pore fluid. Research has shown that many of the mechanical properties of clay soils are closely related to the water-retention characteristics of the soil mass; and since the latter is a function of the mineralogy, pore fluid chemistry, fabric and structure of the soil mass; any change of the pore fluid chemistry may lead to a removal of the existing bonds and therefore an alteration of the internal equilibrium stability of the soil mass (Raymond et al. 1984).

The current approach to landfill liner design also bypasses an accurate prediction of the rate of water movement through the liner under unsaturated conditions and the assumption that clay liners are saturated for the duration of their life is a simplification of

the true process at work. The fact that clay liners are compacted to optimum moisture content, or just wet of optimum, shows that they are initially subject to unsaturated conditions and it is well known that liners can remain variably saturated for the duration of their lives.

### 1.3 Governing Equation

The percolation of moisture through covers and leakage through bottom liners is generally accepted to be governed by the nonlinear Richards Equation (1931) for variably saturated flow:

$$C(\psi) \frac{\partial \psi}{\partial t} = \frac{\partial}{\partial z} \left( K(\psi) \left( \frac{\partial \psi}{\partial z} + 1 \right) \right)$$

This is the “ $\psi$ -based” form of the equation where:  $t$  = time;  $Z$  = elevation;  $K(\ )$  = unsaturated hydraulic conductivity as a function of matric suction; and  $C(\ )$  = specific water capacity (i.e. slope of the soil-water characteristic curve, SWCC). Solving the above Equation requires knowledge of the SWCC and unsaturated hydraulic conductivity.

## 2 Background Information

### 2.1 Soil Theory

A complete review of the basic theory of how energy states are commonly defined in vadose zones is given by T. Karvonen. Soil can be analogized to a bundle of capillaries where the capillary rise is dependent on the diameter of the pore orifice. In reality soils are not simple capillaries; instead water may also move through the soil by forces driven by thermal, electrical, or solute concentration gradients. Water tends to move from higher energy states to lower ones. The total soil water potential is therefore the sum of several forces:

$$H = h_c + h_g + h_{os}$$

where  $h_c$  is soil matric potential (or capillary potential) arising from interactions between soil and water,  $h_g$  is gravitational potential arising from the gravitational force and  $h_{os}$  is

osmotic potential in soil arising from the osmotic forces. The potential can be expressed per unit weight, per unit volume or per unit mass. If the potential is expressed per unit weight, the unit is meters (or cm) ( $\text{J N}^{-1} = \text{m}$ ). When taken per unit weight, the gravity potential corresponds to the elevation head and the capillary or matric potential corresponds to the soil water pressure head ( $h$ ):

$$h_c = \rho_w g h$$

where  $\rho_w$  is density of water (units of mass per volume) and  $h$  is the soil water pressure head (units of length). The soil water pressure head  $h$  is positive below the water table and negative above it. In unsaturated soils, matric potential is negative because tensile forces act on the pore fluid to oppose the matric suction therefore causing a negative net pressure. Matric suction ( $\psi$ ) is the negative of matric potential and is therefore expressed as a positive number.

The term hydraulic head,  $H$ , is sometimes used when dealing only with the sum of matric and gravitational potential and expressed in units of height (or units of length):

$$H = h + h_g = h + z$$

where  $z$  is the vertical coordinate of the system, i.e. height relative to some reference level. Different laboratory methods such as the pressure plate extractor, psychrometer method, and filter paper technique are available to measure components of total suction, matric suction, and osmotic suction. The Pressure plate extractor is a useful tool utilized to sustain specific matric suction on the soil until equilibrium is reached, at which point the moisture content can be measured. From this the SWCC can be derived.

### **2.1.1 Flow through a Soil Element.**

This research will use a computer program which solves the highly nonlinear Richards Equation using a finite difference approach. In order to understand the basic assumptions in this process it is first necessary to describe some basic relationships related to flow through an element.

When considering soil properties it is useful to understand basic nomenclature pertaining to soil characteristics.

- Heterogeneous/Homogenous In a heterogeneous soil the property of interest changes from one point to another throughout the soil (heterogeneous), but stays constant at each point with time (steady state).
- Anisotropic/Isotropic. In an isotropic soil the property of interest is the same in the all directions (x,y,z planes).

### **Steady State Flow**

Steady State, Nonlinear, one-dimensional flow through a finite element can be described as:

$$k_{wy} \frac{d^2 h_w}{dy^2} + \frac{dk_{wy}}{dy} \frac{dh_w}{dy} = 0$$

Nonlinearity comes from second term where coefficient of permeability changes with respect to y-direction due to changing matric suction.

### **Unsteady State Flow**

Richards Equation (1931) is an example of an unsteady state, nonlinear, isotropic relationship. Fredlund and Rahardjo explain how this relationship, based on the Darcy - Buckingham (1907), is derived from a one dimensional net flux of water through a soil element computed from the volume of water entering and leaving the element in a given time:

$$\frac{\partial V_w}{\partial t} = \left( v_{wy} + \frac{\partial v_{wy}}{\partial y} dy \right) dx dz - v_{wy} dx dz$$

Per unit volume this becomes:

$$\frac{\partial V_w}{\partial t} = \frac{\partial v_{wy}}{\partial y}$$

Expressing water volume change as volumetric water content:

$$\frac{\partial(V_w / V_o)}{\partial t} = \frac{\partial v_{wy}}{\partial y}$$

where  $\theta$  is volumetric water content and where:

$$\frac{\partial(V_w / V_o)}{\partial t} = \frac{\partial \theta}{\partial t}$$

and Darcy's Equation:

$$v = k_{wi} \frac{\partial H_w}{\partial y}$$

where  $H_w$  = total head potential

then,

$$\frac{\partial \theta}{\partial t} = \frac{\partial \left( k_{wy} \frac{\partial H_w}{\partial y} \right)}{\partial y} = \frac{\partial}{\partial y} \left( k_{wy} \frac{\partial H_w}{\partial y} \right)$$

Therefore;

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial y} \left( k_{wy} \frac{\partial H_w}{\partial y} \right)$$

or using  $\frac{\partial \theta}{\partial t} = \frac{\partial \theta}{\partial H} \frac{\partial H}{\partial t}$  and  $\frac{\partial \theta}{\partial H} = C(H)$ , specific water capacity:

$$C(H) \frac{\partial H}{\partial t} = \frac{\partial}{\partial y} \left( k_{wy} \frac{\partial H_w}{\partial y} \right)$$

or expressing total head,  $H_w$ , as the sum of gravitational and matric-potentials

$$\frac{\partial H_w}{\partial y} = \frac{\partial(h_w + z)}{\partial y} = \frac{\partial h_w}{\partial y} + 1$$

The above relationship then becomes:

$$C(H) \frac{\partial H}{\partial t} = \frac{\partial}{\partial y} \left( k_{wy} \left( \frac{\partial H_w}{\partial y} + 1 \right) \right)$$

where  $k_{wy}$  is coefficient of permeability as a function of matric suction, otherwise expressed as  $k_y(\psi)$  and where  $C(h)$  is the specific water capacity and is the slope of the -SWCC (volumetric water content vs. matric suction). This is the *h-based* form of Richard's Equation. It is important to note that there are other forms of expressing Richard's Equation including the "θ-based", and the "the mixed-form".

The θ-based:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial y} \left( k \frac{\partial h}{\partial \theta} \frac{\partial \theta}{\partial y} \right) + \frac{\partial k_{wy}}{\partial y}$$

where  $k \frac{\partial h}{\partial \theta}$  is the soil-water diffusivity and is the slope of the soil water retention curve (volumetric water content vs. negative pressure head)

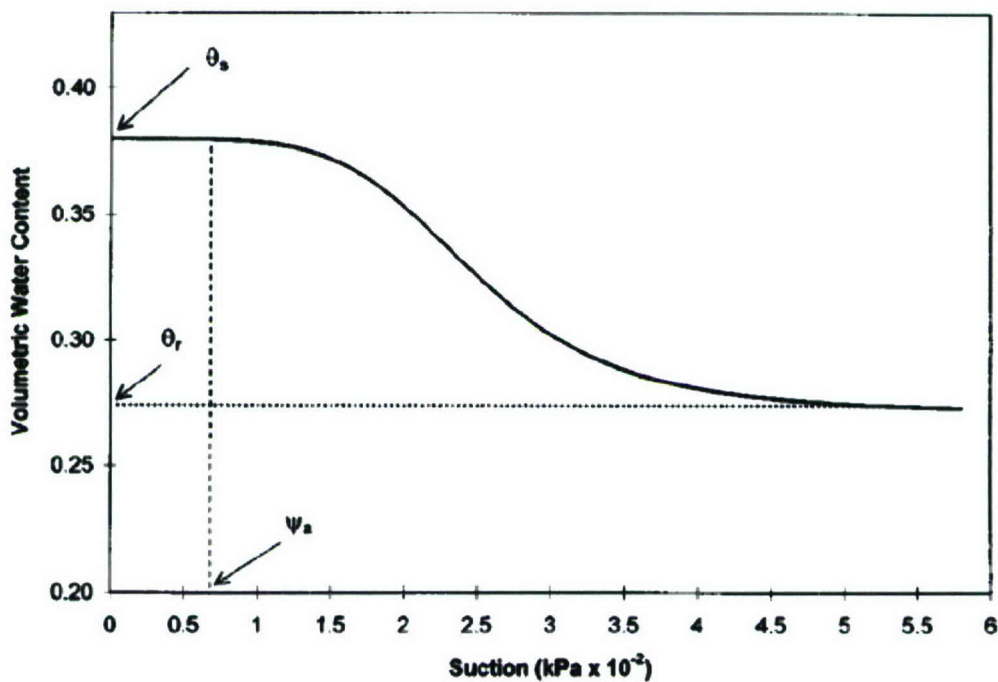
Mixed Form:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial y} \left( k \left( \frac{\partial h}{\partial y} + 1 \right) \right)$$

### 2.1.2 Introduction to the Soil Water Characteristic Curve

Characterizing flow through unsaturated soil requires knowledge of the soil-water characteristic curve (SWCC) and unsaturated permeability. SWCC describes the relationship between matric suction and volumetric water content. As a soil dries, matric suction increases. This is because of capillary forces at work within the porous spaces and to a degree, the hydrophilic chemical nature of clay. The shape of the SWCC depends on soil type, density, soil state, pore size distribution, salt concentrations, temperature, etc (Tinjum et al.). Based on the SWCC, permeability functions, shear strength parameters, and volume change can be predicted. Different laboratory methods such as the pressure plate extractor, psychrometer method, and filter paper technique are available to measure components of total suction, matric suction, and osmotic suction to develop the SWCC relationship. A typical curve that describes the relationship between water content and pore water suction for a clay soil is presented in Figure 1. Several defining parameters of the SWCC are shown, including air-entry suction head ( $\psi_a$ ), residual water content ( $\theta_r$ ), and saturated water content ( $\theta_s$ ). Soils with larger particle

sizes, including sands and silts, would develop a SWCC which plots to the left of the curve shown in Figure 1, with a generally smaller air-entry suction head, smaller residual water content, and smaller value of the saturated water content compared to the curve in Figure 1 (Merayyan 2002). It is well known the SWCC is hysteretic, with bounding curves defining the sorption wetting and desorption drying processes. However, standard practice is to determine only the desorption curve due to experimental difficulties associated with measurement of the sorption curve [Hillel 1980 as discussed in Merayyan et al. (2002)].



**Figure 1: Common Soil-Water Characteristic Curve (from Merayyan et. al “Soil Type...”)**

Soil water characteristic is an important tool in landfill cover and liner analysis and can be used to indirectly measure unsaturated hydraulic conductivity. It can also be used to predict clay cover-liner desiccation cracking, wilting point, and field capacity. The shape of the curve is a function of soil type. Soils with smaller pores tend to have higher air-entry values, while soils with a larger range of pore size tend to exhibit a greater change in matric suction with water content—the slope of the SWCC is steeper (Merayyan 2002).

### **Specific Water Capacity**

A modulus defining the amount of water stored in the soil is required when performing unsteady state or transient seepage analysis. For unsteady state seepage analysis in unsaturated soils, it is the slope of the SWCC which defines the amount of water stored in the soil (Leong and Rahardjo 1993).

### **$\Psi_a$ , $\theta_s$ , and $\theta_r$**

The air-entry value ( $\Psi_a$ ) or the bubbling pressure is the suction required to remove water from the largest pores. This value corresponds to the horizontal portion of the SWCC before the break along the drying portion of the curve. The saturated water content ( $\theta_s$ ) is the water content at which all pores are filled with fluid and the residual water ( $\theta_r$ ) content forms the asymptote approached by the SWCC at very high matric potentials.

#### **2.1.3 SWCC Functions**

Soil water characteristic functions provide mathematical representations of the soil water characteristic curve based on empirical parameters. These functions vary in their ability to describe the curves and their applicability may change depending on the soil properties. This study considered two different characteristic functions.

The Brooks Corey:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{\left[ (h/\Psi_a)^\lambda \right]}$$

The Van Genuchten:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{\left[ 1 + (-\alpha h)^n \right]^m}$$

Both of these models require parameter optimization in order to uniquely describe a soil. Brooks Corey requires optimization of  $\theta_r$ ,  $\Psi_a$ , and  $\lambda$ .  $\lambda$  is the pore-size distribution index in which Brooks Corey is heavily dependent on. Through Kelvin's capillary law, the pore-size distribution index is assumed to be analogous with the slope of the soil-water characteristic curve (Leong and Rahardjo 1997).

The Van Genuchten model is defined by three parameters,  $\alpha$ ,  $n$ , and  $m$ . Each of these parameters is described by Leong and Rahardjo (1997). The parameter  $\alpha$  is the pivot point of the curve, and its value is inversely related to the air-entry suction (Tinjin et al.) As  $\alpha$  increases, the air-entry suction decreases. The parameter  $n$  controls the slope of the SWCC about the pivot point, which occurs at a normalized volumetric water content ( $\Theta$ ) of 0.5, where:  $\Theta = (\theta_w - \theta_r) / (\theta_s - \theta_r)$ . As  $n$  increases, the sloping portion of the curve between  $\psi_a$  and the knee (the point of inflection at the lower portion of the curve as it approaches a horizontal position) of the SWCC becomes steeper (Merayyan 2002). The parameter  $m$  rotates the sloping portion of the curve. As  $m$  increases, the range of the curve between  $\psi_a$  and the knee of the SWCC decreases. The stability of the curve-fitting process is improved by equating the parameter  $m$  to  $1 - (1/n)$  (Van Genuchten et al. 1997). This relationship was adopted for this study.

#### 2.1.4 Permeability Functions

Permeability functions relate unsaturated permeability to matric suction. Since matric suction is related to volumetric water content by the SWCC, permeability can also be derived as a function of volumetric water content. Permeability functions can be derived through indirect or direct methods utilizing a steady or non-steady state approach (Leong and Rahadjo 1997)

Direct methods require a rigorous data collection of permeability at various matric suctions. A steady state method might involve imposing matric suction on the soil using the axis translation technique, and once equilibrium is reached, denoted by constant water content, the permeability is measured. An unsteady state method would involve imposing a continuous flow of water through a cylindrical specimen, and then measuring the hydraulic gradient and matric suction at various points along the flow path (Leong and Rahadjo 1997).

In indirect methods, permeability is inferred from the soil water characteristic curve. Like SWCC functions, they provide a mathematical relationship which will vary in their applicability to unique soils. Indirect methods can be further divided into two categories—empirical and statistical models.

Empirical equations for unsaturated hydraulic conductivity are a function of volumetric water content ( $\theta_w$ ) or matric suction ( $\psi$ ). These functions require parameters from laboratory data. The Van Genuchten model is an example of this type of conductivity function:

$$k(h) = k_s \frac{\left\{ 1 - (-\alpha h)^{n-1} \left[ 1 + (\alpha h)^n \right]^{-m} \right\}^2}{\left[ 1 + (-\alpha h)^n \right]^{m/2}}$$

where  $\alpha$ ,  $n$ , and  $m$  are empirical constants. By using an indirect, empirical permeability function, a comparison of unsaturated fluid movement based on the soil water characteristic can be made. This was the approach adopted in this research. The Van Genuchten model for unsaturated permeability was also selected to provide continuity with the Van Genuchten SWCC function.

Based on a critical analysis of various empirical conductivity functions, Leong and Rahardjo (1997) have concluded that research on empirical conductivity functions have suggested the generalized relationship:

$$k_r = \Theta^p$$

Where  $k_r$  is relative coefficient of permeability or ratio of permeability  $k_w$  and saturated permeability  $k_s$ ; and  $\Theta$  = normalized volumetric water content or  $\Theta = (\theta_w - \theta_r) / (\theta_s - \theta_r)$  where subscripts  $r$  and  $s$  denote residual and saturated respectively; and  $p$  is a constant. Leong and Rahardjo (1997) concluded that this relationship fits the data reasonably well. Therefore if  $p$  is known for a particular soil, then permeability can be indirectly determined from the soil-water characteristic curve. The significance of Leong and Rahardjo's  $k_r = \Theta^p$  relationship in this research is that reasonable conclusions can be made about the impacts of different permeates on unsaturated conductivity by assessing changing SWCC parameters only. If  $p$  is an experimental control (i.e. the soil is not changing), then changes in  $k_r$  are a function of normalized volumetric water content only.

Statistical permeability functions derive permeability coefficients from the SWCC. In these determinations, the relationship is derived based on Poiseuille's Equation (Leong and Rahardjo 1997) and is heavily reliant on a statistical pore size distribution index,  $\lambda$ .  $\lambda$  is derived based on the assumption that pore size distribution is analogous with the

SWCC (Leong and Rahardjo 1997). The Brooks Corey permeability function is an example of a statistical permeability function.

$$k(h) = k_s \frac{k_s}{[h/h_b]^n}$$

Another very common statistical permeability function is Campbell's:

$$k_u = k_s \left[ \frac{\theta - \theta_r}{\phi - \theta_r} \right]^{3 + \left( \frac{2}{\lambda} \right)}$$

### 3 Literature Review

#### 3.1 Properties of Clay that may be Influenced by Leachate

The electrical double-layer is an important component to the internal stability of clayey soils. Clays typically fall into a size range appropriate for colloids. Colloids are particles that are sufficiently small enough to allow electrostatic or interfacial forces to significantly impact their behavior. For clay minerals, these forces are present as a charged solid-solution interface that forms an electric or diffused double layer (Goldman et al. 1990). The thickness of this layer is an important factor in engineering properties of clay. In general, increasing the clay-mineral electrical double layer thickness will result in a more dispersed fabric and lower hydraulic conductivity (Goldman et al. 1990). Mitchell and Madsen concluded that any changes in hydraulic properties of the clay, such as permeability, were most likely the result of changes to the soil fabric. Because the electric double layer is highly dependent on the electro-chemical properties of the solution-solid interface, any change in the pore fluid chemistry could impact the layer and thus change the soil fabric.

Goldman et. al (1990) have neatly summarized various studies on clay-chemical interactions that influence permeability. They state that changes in permeability of clay soils due to chemical interactions may result from the following processes:

1. Alterations in soil fabric stemming from chemical influences on the diffuse double layer surrounding clay particles.
2. Dissolution of soil constituents by strong acids or bases.

3. Precipitation of solids in soil pores.
4. Soil pore blockage due to the growth of microorganisms.

### **3.2 Recent Investigations**

Extensive research has been done to analyze the physiochemical changes of clay when subjected to organic and inorganic substances. However research in this area has lacked any comprehensive conclusions and can be contradictory. Numerous publications were produced in the mid to late 1980's on soil-fluid compatibility especially relating to organic and inorganic effects on intrinsic permeability.

A literature review completed by H.Y. Fang (1997) concluded the following:

1. Hydraulic conductivity increases as the percentages of multivalent cations in the pore fluid increases;
2. Hydraulic conductivity increases when organics are introduced; these increases may be related to the dipole moment and dielectric constants of the solvents.
3. Hydraulic conductivity increases as the salt concentration of pore fluid increases.

Fang and Evans (1988) observed virtually no difference in the saturated permeability of clay samples permeated with water and landfill leachate. Mitchell and Madsen (1987) summarized literature on the effect of inorganic and organic on saturated clay permeability. The summary indicates that although dilute solutions of inorganic chemicals may change clay permeability, dilute solutions of organic chemicals have virtually no effect. They also found significant changes in permeability of the clay samples were not observed when leached with salt solutions.

It was concluded that clay-chemical interactions are generally more pronounced in the case of high water content systems because particles are less constrained from movement which might occur due to a change in soil fabric. Mitchell and Madsen also found that permeability was highly dependent on the testing method.

Permeability testing by Mitchell & Madsen and Fang & Evans used either a flexible wall permeameter for saturated samples or a rigid wall permeameter test for compacted soils to determine the hydraulic conductivity. These methods are fundamentally different from an unsteady-state approach to determining unsaturated hydraulic conductivity. Unsaturated permeability which is based on pore distribution and matric suction could more accurately describe the system in regard to soil fabric, thus giving a better indication of whether leachate has an affect on flow through clay landfill liners.

Most Recent research by Ichola & Gaidi (May, 2006) found an increase in hydraulic conductivity when leachate is used as pore fluid as compared to water — however, the increase was not more than  $1.2 \times 10^{-10}$  m/s. Infiltration tests were carried out on a clay soil column and various methods were employed to determine water retention curves, including Richards pressure plate apparatus, filter paper method, and the equilibrium relative moisture method. A Van Genuchten model was used to describe the water retention curve and Brooks Corey (1964) to compute hydraulic conductivity as a function of water content and matric suction.

## **4 Clay and Leachate Descriptions**

### **4.1 Leachate**

Leachate was sampled from the City of Paso Robles Landfill located outside of Paso Robles, CA in San Luis Obispo County. This is a class III waste management facility permitted by the Regional Water Quality Control Board to dispose non-hazardous municipal solid waste. Typical waste types received for disposal are summarized in Table 1.

**Table 1: Waste Constituents of Landfill where Leachate was Sampled**

Waste Constituent	Percent
Residential Curbside Collection Wastes	21.5%
Comercial Bin Service Discards	40.0%
Roll-Off Box Debris	5.9%
Self-Hauled Wastes	14.9%
City of Paso Robles Waste Debris	2.7%
Sewage Treament Plant Grit Screenings	0.4%
Mattresses and Sofa's	0.2%

Just less than 5.0 gallons of leachate was sampled from the leachate system that provided recirculation to a closed portion of the landfill. The sample was collected in a 5.0 Gallon HDPE container and was immediately sealed with an air-tight lid. The sample was kept at room temperature (approximately 60°F) and sheltered from any direct natural or artificial light for the duration of laboratory testing. It was beyond the scope of this research to provide analysis of leachate constituents; however the March 2003 Detection Monitoring report submitted to the California Regional Water Quality Control Board revealed the following leachate characteristics:

**Table 2: Possible Leachate Constituents**

Analyte	Result	Units
Total Dissolved solids	1240	mg/L
Chloride	294	mg/L
Sulfate	91.9	mg/L
Sodium	246	mg/L
Barium	0.204	mg/L
Copper	0.014	mg/L
Lead	0.026	mg/L

No volatile organic compounds (VOCs) were detected in the 2003 analysis; however, past reports have indicated the presence of organic compounds:

**Table 3: Possible Leachate VOC Constituents**

Analyte	Result	Units	Date Sampled
Chloroethane	42	µg/L	January, 1995

Analyte	Result	Units	Date Sampled
1,1-Dichloroethane	64	µg/L	January, 1995
Methylene Chloride	350	µg/L	January, 1995
Tertachloroethane	29	µg/L	January, 1995
Toluene	23	µg/L	January, 1995
1,1,1-Trichloroethane	41	µg/L	January, 1995
Trichlorofluoromethane	24	µg/L	January, 1995
Xylenes	14	µg/L	January, 1995

## 4.2 Clay

The clay used for this research was sampled from a stockpile which is currently supplying low permeability clay for the construction of two landfill liners for the Chicago Grade Landfill in the Atascadero-Templeton area of Northern San Luis Obispo County, California. The landfill is owned and operated by the Chicago Grade Landfill, Inc. and consists of approximately forty-five acres of permitted land. Locally available clays have been used in the construction of other onsite Subtitle D liners. This clay is derived from the Paso Robles Formation, which consists of poorly indurated, thinly bedded to massive sediments ranging from clay to fine gravel. The material which has been determined to be low permeability has been segregated and stockpiled.

Clay was sampled from a clay quarry at the Chicago Grade Landfill. The clay was excavated and immediately placed in plastic containers. The containers provided adequate moisture insulation, but they were not air tight. Clod size ranged from a couple millimeters in diameter to, roughly 75mm in diameter. Field moisture content was later determined to be about 29 to 30%.

### 4.2.1 Geotechnical Investigations

Laboratory and field hydraulic conductivity tests by Hoover and Associates have consistently indicated hydraulic conductivity measurements in the  $1 \times 10^{-8}$  to  $1 \times 10^{-7}$  cm/sec range when compacted to 90% relative compaction on the wet side of optimum. The geotechnical report by Hoover & Associates also revealed the following properties:

**Table 4: Previous Geotechnical Investigation**

Soil Type	Liquid Limit	Plastic Index	Scaled Double Ring Infiltrometer
-----------	--------------	---------------	----------------------------------

Chicago Grade Brown Matrix	40	14	9.54x10 <sup>-8</sup> to 9.12x10 <sup>-7</sup> cm/s
----------------------------	----	----	---

Part of this research included a completed geotechnical analysis. This investigation was conducted to determine two things:

1. The similarity between published properties of the clay used for liner construction.
2. Pertinent information to this research such as specific gravity, maximum dry density, optimum moisture content, and Plasticity Index.

The analysis in this research investigated the following soil properties:

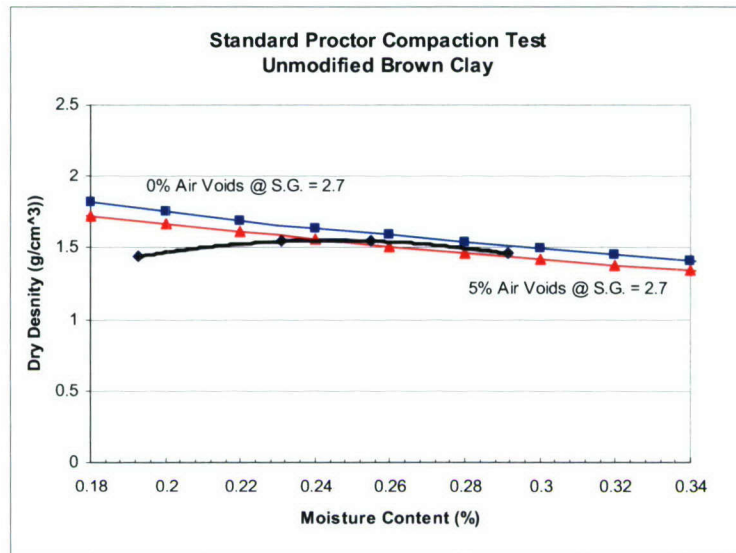
1. Soil Compaction testing for maximum dry density and optimum water content
2. Atterberg Limits
3. Specific Gravity of solids
4. Residual Moisture Content
5. USCS Classification and Grain Size Distribution

### **Soil Compaction Testing**

Compaction testing was carried out in accordance with ASTM D698. This analysis yielded the results shown in Table 5 and in Figure 2.

Table 5: Standard Compaction Results

Soil Type	Standard Proctor Maximum Dry Density	Optimum Moisture Content
Chicago Grade Soil	96.6 PCF	25.5 %



**Figure 2: Standard Proctor Compaction Test Results**

### Atterberg Limits

The Atterberg limits were carried out in accordance with ASTM D 4318 and the results are summarized in Table 6.

**Table 6: Atterberg Limits**

Soil Type	Liquid Limit	Plastic Index
Brown Clay Sample	39	14

Laboratory testing revealed the clay sample to have practically the same plasticity index as tested by Hoover & Associates. The clay is therefore representative of the clay used in the Chicago Grade Landfill liner construction.

### Specific Gravity

The specific gravity of clay solids was carried out as specified in ASTM C 127, ASTM D 70, and ASTM D 854s. The table below summarizes the results.

**Table 7: Specific Gravity**

Soil Type	Specific Gravity
Brown Clay Sample	2.68

### **Residual Moisture Content**

The residual water contents were determined subsequent to an extended period of air drying. The samples were allowed to air dry in the laboratory under controlled temperature, approximately (65° F) for a period of three weeks. The samples were weighed and oven-dried for 24 h to determine residual gravimetric water content. The results for this analysis can be found in the Section 6.2.

### **USCS Classification and Grain Size Distribution**

These properties are shown in Table 8.

**Table 8: USCS Classification and Grain Size Distribution**

Soil Type	USCS Classification	Grain Size Distribution			
		Gravel	Sand	Silt	Clay
Brown Clay Sample	Sandy-Lean Clay (CL)	8%	28%	13%	51%

i.

## **5 Introduction to Numerical Model and Pressure Plate Extractor**

### **5.1 Pressure Plate Extractor**

For this study the axis translation method was used. This is a simple and effective way to measure soil water characteristic curves.

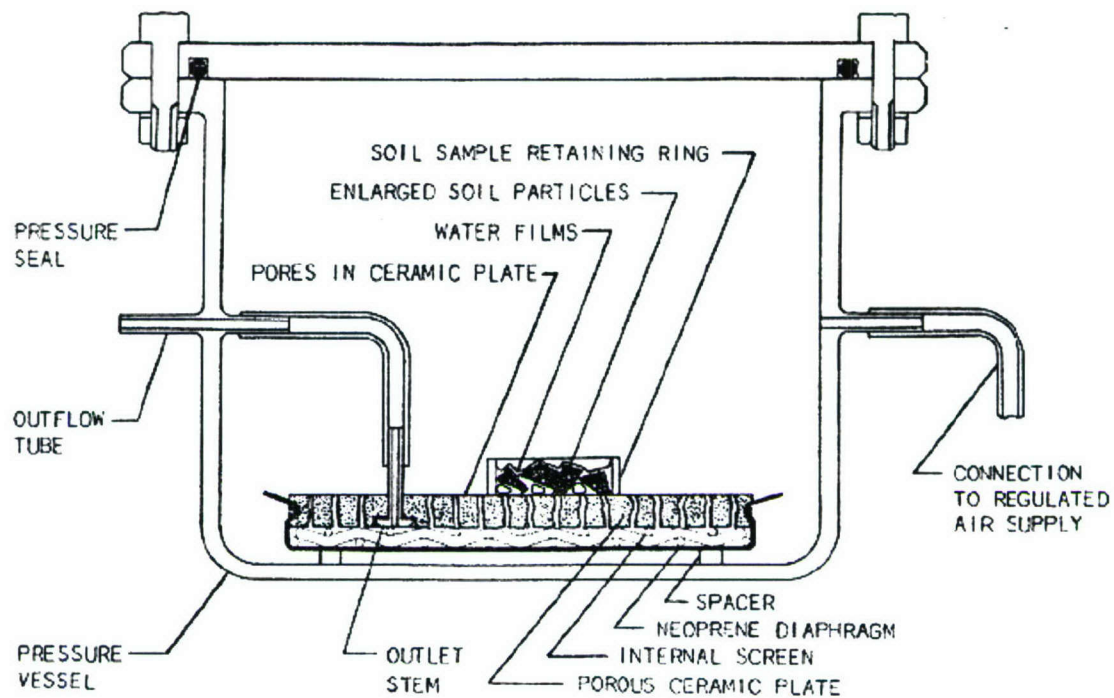
Hillel (1956 as discussed in Merayyan et al. (2002)] demonstrated that the pore-water pressure is increased by an amount equal to the increase in the ambient chamber air-pressure. In other words, the soil matric suction remained constant when measured at various ambient pressures.

A Pressure Plate Extractor (PPE) is used to provide ambient pressures above atmospheric. The pressure tank, valves, gages, and ceramic plates were all obtained from Soil Moisture Corporation, in Santa Barbara, a picture of which is provided in Figure 3 while diagrams of the PPE system is given in Figure 4 and Figure 5.



**Figure 3: Pressure Plate Extractor and Pressure Regulator**

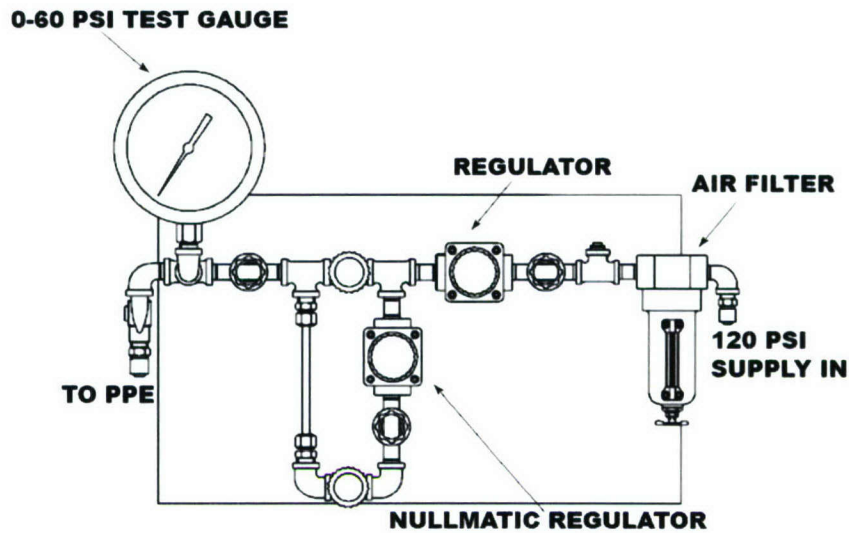
The PPE is capable of imposing matric suctions up to 5.0 Bars. Four ceramic plates, capable of withstanding air-entry values up to 0.5, 1.0, 3.0, and 5.0 Bars are used in this research. The ceramic plates are manufactured with microscopic pores to allow the passage of water. As long as the pressure in the extractor does not go above the air entry value of the plate, air will not flow through the pores because of surface tension forces at the gas-liquid interface. The air-entry value is a function of the pore diameter (Soil Moisture Corp 2006).



**Figure 4: Diagram of the Pressure Plate Extractor**

When the soil is placed on the ceramic plates, and the extractor is pressurized, water will begin to flow from around each of the soil particles and out through the ceramic plate. Equilibrium occurs when the effective curvature of the water films throughout the soil are the same as the pores in the plate (Soil Moisture Corp 2006).

Pressure is supplied to the PPE through an available 120 PSI supply line which is controlled with one standard regulator in series with a nullmatic regulator. Together, the two regulators can control a stable pressure supply within 0.5 PSI (0.02BAR) accuracy. The nullmatic regulator has a maximum capacity of 4.0 bars and must be bypassed beyond this range. The main difference in using this bypass was an increased difficulty in setting the target pressure. Once the target pressure was set, no fluctuations were visible.



**Figure 5: Diagram of the PPE Pressure Regulation System**

## 5.2 Numerical Analysis Description

CHEMFLO-2000 is a free software available through the EPA which enables users to simulate water movement and chemical fate and transport in vadose zones. Water movement and chemical transport are modeled using Richards Equation, and the convection-dispersion equations, respectively. The equations are solved numerically using a finite differences approach. CHEMFLO-2000 is an upgraded version of CHEMFLO V1.3 that was released in 1989.

CHEMFLO uses Richards Equation to describe one-dimensional water movement in the following form:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial y} \left( K(H) \left( \frac{\partial h}{\partial y} - \sin(A) \right) \right)$$

where  $\theta = \theta(h)$  is the volumetric water content;  $h=h(y,t)$  is the matric potential and  $x$  is position coordinate parallel to the direction of flow;  $t$  is the time;  $\sin(A)$  is the sine of the angle  $A$  between the direction of flow and the horizontal direction; and  $K(h)$  is the hydraulic conductivity of the soil at matric potential  $h$  (Nofziger and Wu 2003).

Appropriate initial and boundary conditions are applied to Richards' equation in order to define the water flow problem. The required soil hydraulic properties are defined by specifying the  $\theta(h)$  and  $K(h)$  functions.

### 5.3 Limitations

In any numerical modeling or experimental investigation study it is important to fully understand limitations of the analysis. The numerical analysis in this research has the following limitations.

1. The numerical model used in this research assumes that flow and transport in the soil is strictly one dimensional. Flow in the field will often be multi-dimensional due to layers within the profile, spatial variability of soil properties, and spatially variable application rates.
2. The Richards Equation for water movement is based on the Darcy-Buckingham equation for water movement in unsaturated soils. This equation is usually a good descriptor of water movement in soils, but exceptions exist. No provision is made in this numerical model for clay swelling or preferential flow of water through large pores in contact with free water. Therefore, it will not accurately represent flow in soils with large cracks.
3. This analysis assumes hysteresis in the wetting and drying processes can be ignored.
4. The simulated results depend upon the initial conditions specified. If the specified initial conditions do not match the real conditions, the calculated values may be incorrect.
5. The predictions in CHEMFLO can be quite sensitive to the specified boundary conditions. If the specified ones do not match the actual conditions, large errors may be made.
6. Limitations in the results due to approximating derivatives by finite differences, as well as other approximations used in solving the partial differential equations,

are subtle and are often difficult to detect (Nofziger and Wu 2003).. Mass balance errors for water and chemicals are calculated in CHEMFLO to detect net computational error. Small mass balance errors are essential conditions for a valid solution, but they do not guarantee accurate solutions. In general, these errors tend to decrease as the mesh sizes decrease.

7. Limitations exist based on resolution of experimental data. The pressure plate extractor is a robust procedure for measuring matric suction as a function of water content; however, limitations exist based on the resolution to which changes in matric suctions can be measured with respect to changes in water content.

### 5.3.1 Soil Water Characteristic and Permeability Inputs to CHEMFLO

CHEMFLO allows parameter input for multiple conductivity and soil water characteristic functions.

**Table 9: CHEMFLO Conductivity and SWCC Functions**

Conductivity Functions		Soil-Water Functions	Characteristic Functions
Gardner	$K(h) = K_s \frac{K_s}{1 + [h/h_c]^n}$	Simmons	$\Theta(h) = \alpha - \ln[(-h)^n]$
Brooks Corey	$K(h) = K_s \frac{K_s}{[h/h_b]^n}$	Brooks Corey	$\Theta(h) = \Theta_r + \frac{\Theta_s - \Theta_r}{[(h/h_b)^\lambda]^2}$
Van Genuchten	$K(h) = K_s \frac{\{1 - (-\alpha h)^{n-1} [1 + (-\alpha h)^n]^m\}^2}{[1 + (-\alpha h)^n]^{m/2}}$	Van Genuchten	$\Theta(h) = \Theta_r + \frac{\Theta_s - \Theta_r}{[1 + (-\alpha h)^n]^m}$

In this research, parameter optimization is used to optimize input parameters to the Van Genuchten and Brooks Corey's soil water characteristic based on data from the Pressure Plate Extractor. It was beyond the scope this project to run a separate parameter optimization for permeability functions because permeability was not measured in the lab. Furthermore, permeability measurements are not necessary because the focus of this

research is to isolate the affects of the soil water characteristic function on fluid movement. The Van Genuchten Conductivity function is selected because its parameter requirements are the same as those in the SWCC function. It is assumed an estimate of the conductivity function could be made by only knowing the parameters of the SWCC—which in this case have been tested and optimized. The precise applicability of the SWCC parameters to the permeability function is inconsequential because they act as a control. It has been noted that any change in SWCC will also have an affect of the permeability function itself. Leong and Rahardjo have concluded that a basic relationship exists between the SWCC and the permeability function:

$$k_r = \Theta^P$$

If  $P$  is known for a particular soil, then permeability as a function of matric potential can be derived. In this case,  $P$  is not known, but because the *change* in permeability is of interest and the soil is the same, then difference in permeability is a function of changes in the SWCC. In the final results it is not so much the absolute values that will be of interest, instead the differences between leachate and water are emphasized.

## **6 Methodology**

Three main methods have been employed in order to achieve the objectives of this research—these methods include Laboratory Testing, Parameter Optimization, and Numerical Analysis in CHEMFLO. These methods are described in this section.

### **6.1 Laboratory Testing**

Laboratory Testing includes preparing the samples for the Pressure Plate Extractor (PPE) and testing the samples using the Axis-Translation technique.

#### **6.1.1 Sample Preparation**

Samples are prepared for the pressure plate extractor (PPE) in plastic retaining rings 1.0 cm high and 5.5 cm in diameter. It is desirable to keep sample heights small in order to keep the time to reach equilibrium reasonable. The samples are removed from a standard compaction mold using a hydraulic jack and carefully cut using a long, sharp knife and trimmed before being placed in the retaining rings. All cutting is done on a solid metal

surface which has been slightly coated with excess soil from the compaction test. This is done to prevent moisture from leaching into the pores of the cutting surface in order to preserve moisture in the edges of the sample. An initial cross-sectional slice is made equal to the height of the retaining rings. The slice is then trimmed to the diameter of the ring, taking extra care not to disturb compaction conditions. The slice is then placed in the retaining ring and set horizontally on the cutting surface. Any soil which protrudes above the top of the ring is trimmed.

To obtain better kneading action and lower permeability in landfills, field compaction is specified to be done at 3 to 4% wet of optimum (A. Bagchi). In order to mimic field conditions, samples are prepared for the pressure plate extractor at 4% wet of optimum using standard compaction. Results from compaction testing yielded a 25.5% optimum moisture content. The target gravimetric water content is therefore 29.5%, with an allowable difference of 5%.

$$\text{Percent Difference} = \frac{|\text{target} - \text{actual}|}{\left(\frac{\text{target} + \text{actual}}{2}\right)} \leq 5\%$$

This percent difference results in an acceptable range of starting moisture contents from 26.5% to 30.8%. All samples are within this range, most of which fall below a 3% difference. It is important to keep a tight range of moisture contents because, as Merayyan et al. have found, compaction conditions and initial moisture content can influence the soil water characteristic curve.

All samples are dried in an oven at 90 degrees Fahrenheit until reaching essentially zero water content and then brought to the proper water content using either tap water or leachate, depending on what permeate is being tested. The dried soil is placed into air-tight ceramic containers large enough to hold enough soil for one standard compaction test. An appropriate amount of fluid is then added to the container to reach the target equilibrium moisture content. The lid to the container is closed and the fluid is allowed to be absorbed for a period of five to seven days. When the soil is ready to be prepared for the PPE, the soil is removed and a microwave method is used to test a small portion of the soil for water content. If this preliminary test reveals that the target water content has

been reached, the soil is then compacted and prepared for the PPE. If not, the appropriate pore water fluid is added using a spray bottle to uniformly add moisture to the specimen, the soil is placed back into the containers and allowed to reach equilibrium for a period of twenty-four hours. No more than 1% gravimetric water content had to be added to any of the samples used in this research. A total of three tins with representative soil from the top, middle, and bottom of both compaction molds are then oven-dried according to ASTM D2216 to determine moisture content. This serves to verify that the compaction moisture content is within acceptable limits.

The field water moisture content of 29% to 30 % was conveniently close to the target moisture content of 29.5%. If the clay was slightly dry before compaction, additional water was uniformly added using a spray bottle. If the samples were slightly wet, the clay was allowed to air dry until the target water content was reached.

### **6.1.2 Laboratory Testing**

#### **Testing Summary**

The purpose of laboratory testing is to produce data points on the soil-water characteristic curve. Discrete points are developed by pairing matric-suction values with a single volumetric moisture content. One “test” can be defined as the procedure which produces this unique point. For each test, twelve samples are carefully cut from the compaction molds, placed into confining rings, and then positioned on a ceramic plate in the PPE. The resulting water contents are measured at four imposed matric suctions—0.5, 1.0, 3.0, and 5.0 BARs. The PPE is initially pressurized to 0.5 BAR until equilibrium is reached. Three samples are then removed from the chamber and oven dried to determine their moisture content. The PPE and the remaining samples are then pressurized to the next incremental pressure, up to 5.0 BARs. This procedure is repeated three times for both leachate and water samples.

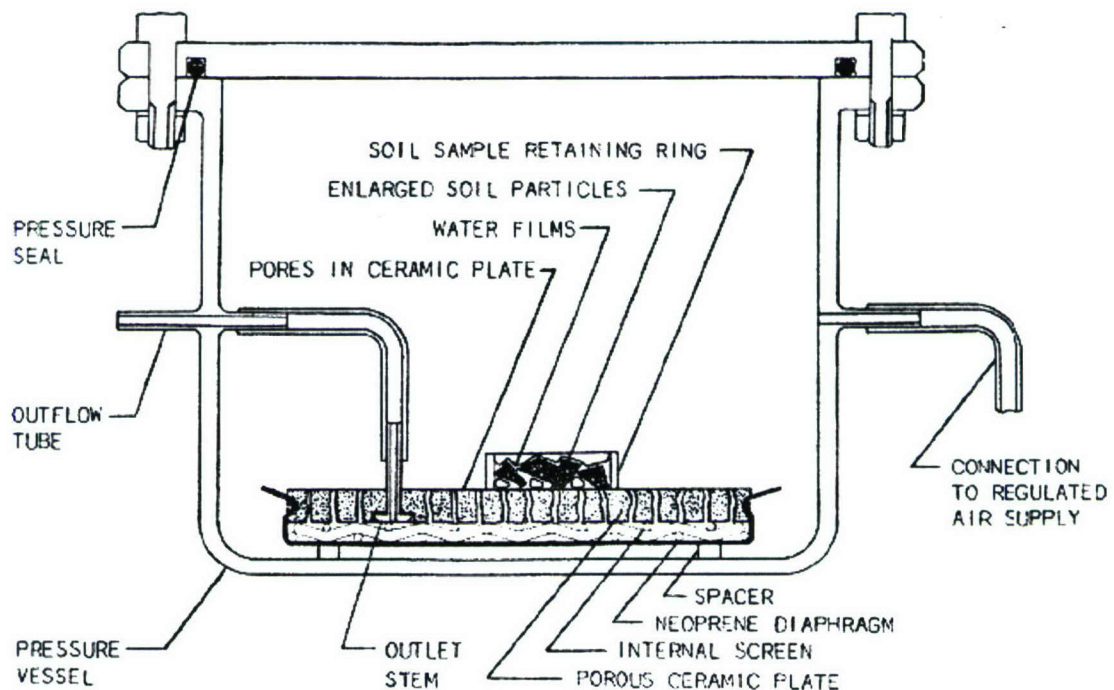
#### **Testing Details**

A total of four ceramic plates are used; each one designed for a unique equilibrium pressure at ½ Bar, 1 Bar, 3 Bar, and 5 Bar. Each plate can accommodate a maximum of twelve sample rings. The plates are submerged in water for a minimum of thirty-six hours to fully saturate the pores. To prepare the plate for testing it is removed from water, held

upside-down, and allowed to free drain via the outlet stem until all excess water in the reservoir is removed.

In the first phase of a test run, the prepared sample rings, as described in section 6.1.1, are placed on the  $\frac{1}{2}$  bar plate. If the excess moisture on the ceramic plate has evaporated, just enough water is sprayed on the surface to produce a visible shine before placing the samples on the plate. This moisture will promote a better connection between ceramic pores and the clay samples. To ensure intimate contact has been made, the rings are slightly pushed on and rotated 90 degrees. Once again, care is taken not to disturb compaction conditions.

The ceramic plate is then placed inside the PPE. The outlet stem is aligned with the outflow and connected via a neoprene tube. The PPE lid is inspected to ensure the seal is free of soil particles and then placed on the PPE and hand-tightened with wing-nuts.



#### Setting the Target Equilibrium Pressure.

1. Initially all valves and regulators are closed.

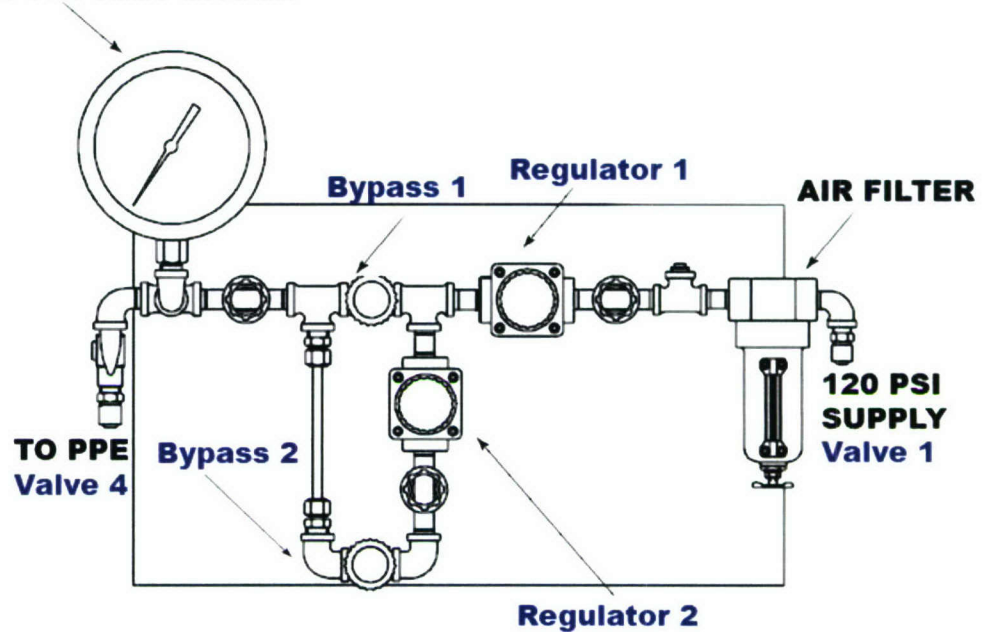
2. The nullmatic regulator (Regulator 2) is open to  $\frac{3}{4}$  full.
3. Valve 1 is then fully opened.
4. When equilibrium is reached—when hissing sound has dissipated—Bypass 2 is fully opened.
5. Regulator 1 is opened until pressure on the Test Gauge reads about two to three PSI above the target pressure.
6. The nullmatic regulator is then slowly closed until the target equilibrium pressure is reached.
7. Finally, Valve 4 is slowly opened until the PPE is fully pressurized and then completely opened.

The nullmatic regulator has a maximum capacity of 4.0 bars, therefore it must be bypassed beyond this range. Only one test was carried out at beyond 4-Bars (5-Bars). The pressurizing procedures are then as follows:

1. Initially all valves and regulators are closed.
2. The nullmatic regulator (Regulator 2) is open to  $\frac{3}{4}$  full.
3. Valve 1 is then fully opened.
4. When equilibrium is reached—when hissing sound has dissipated—Bypass 1 is fully opened, and Bypass 2 remains closed
5. Regulator 1 is opened until pressure on the Test Gauge reaches the target pressure.
6. Finally, Valve 4 is slowly opened until the PPE is fully pressurized and then completely opened.

7. Regulator 1 is then adjusted again to reach desired equilibrium pressure.

**0-60 PSI TEST GAUGE**



The outflow tube is connected to a graduated cylinder with a graduated measurement of 0.01ml. The tube connects to a plastic cap which sits on the inside rim of the graduated cylinder. The cap minimizes evaporation while still exposing the cylinder to the atmosphere. Once all the samples are in the PPE and the pressure has been set, the PPE is allowed to reach equilibrium. Equilibrium is assumed when fluid flow is less than 0.01ml/day.

When the equilibrium is reached, the outflow tube is pinched off and the samples are removed from the PPE using the following procedure:

1. Valve 1 is closed.
2. Regulator 2 is fully closed. This releases pressure in the PPE.
3. When all pressure has been released, the lid is removed.
4. Three sample rings are then removed to determine the equilibrium moisture content. To eliminate local variations in the plate's ability to transmit moisture away from the sample, samples at each equilibrium pressure are selected from

different locations around the plate. These samples are placed in tin containers and allowed to oven dry according to ASTM D2216 to determine moisture content.

5. The next incremental ceramic plate is prepared like before and the remaining samples are placed on it also in the same manner.

The procedure is repeated for the remained target equilibrium pressures.

## 6.2 Parameter Optimization Methods

The Brooks Corey and Van Genuchten soil-water characteristic models were chosen to model the soil water characteristic curves, and thus it is these models for which parameter optimization was conducted. Data from the initial geotechnical investigation, the sample preparation information, and the experimental PPE data are used to optimize the parameters in these models. Table 10 describes all the necessary input data, to both models, and its source as it applies to parameter optimization.

**Table 10: Input Data to Parameter Optimization**

Data	Symbol/Units	How Data is utilized	Data Source
Specific Gravity	$\rho_s$	Calculating Phase relationships	Geotechnical Investigation--ASTM C 127, ASTM D 70, ASTM D 854,
Density of Water	$\rho_w (\text{g/cm}^3)$	Calculating Phase relationships	Assumed at 20° Celsius
Dry Density	$\rho_d (\text{g/cm}^3)$	Calculating Phase relationships	Sample Preparation measurement per ASTM D 698
Saturated Volumetric Water Content	$\Theta_s$	Necessary to model SWCC	Calculated from Phase Relations
Residual Volumetric Water Content	$\Theta_r$	Necessary to model SWCC	Geotechnical Investigation
Matric Suction	$\psi (\text{cm})$	Necessary to Compute theoretical volumetric Water content	PPE Testing
Measured Volumetric Water Content	$\Theta_m$	Necessary to optimize parameters	PPE Testing
Computed Volumetric Water Content	$\Theta_c$	Necessary to optimize parameters	Calculated from Van Genuchten SWCC Function

Data	Symbol/Units	How Data is utilized	Data Source
Brooks-Corey Parameters	$h_b$ (cm)	Necessary to compute theoretical volumetric water content	Optimized Parameters
	$\lambda$		
Van Genuchten Parameters	$\alpha$ (cm) or (1/cm)	Necessary to compute theoretical volumetric water content	Optimized Parameters
	$n$		
	$m$		
Sum of Squared Residuals	SSR	Parameter Optimization Process	Parameter Optimization Process

The resulting data from the Pressure Plate extractor data was in the form of equilibrium gravimetric water contents (WC) and their respective equilibrium pressures. The Brooks Corey and Van Genuchten Soil water characteristic functions model the SWCC in terms of volumetric water contents ( $\Theta$ ) —it was therefore necessary to derive volumetric water contents from the gravimetric water contents phase relations. The following describes the phase relationships used to develop the necessary parameter optimization data:

Volumetric Water Content ( $\Theta$ ):

$$\Theta = (WC)(\rho_d \rho_w)$$

It is also necessary to compute Saturated Water content from phase relations.

Saturated Volumetric Water Content ( $\Theta_s$ )

$$\Theta_s = (WC_s)(\rho_d \rho_w)$$

$$WC_s = (S)(\rho_s \rho_w - \rho_d \rho_w) / (\rho_s \rho_d)$$

The soil is saturated, so the degree of saturation is equal to 1.

$$\Theta_s = [(1)(\rho_s \rho_w - \rho_d \rho_w) / (\rho_s \rho_d)] * \rho_d \rho_w$$

$$\Theta_s = (\rho_s \rho_w - \rho_d \rho_w) \rho_w / (\rho_s)$$

Solver was used to minimize the sum of the squared residuals calculated as the difference between the measured and computed volumetric water contents. The sum of the square residuals can be calculated as the following:

$$SSR = \sum_{i=1}^n w_i (\theta_{mi} - \theta_{ci})^2$$

Where  $w_i$  is a weighting factor;  $\theta_m$  measured water content at a certain pressure level; and  $\theta_c$  is the calculated water content from each model at the same pressure level. The weighting factor was set equal to 1.0 as all data points were assumed to be equally significant. The best fit of each model to the measured data was assumed to be the one that resulted in the minimum SSR value.

### 6.2.1 Solver

Soil Water Characteristic parameter optimization was conducted using Microsoft Excel's Solver routine. Solver is part of a suite of commands which operated within Excel's programming structure. Solver allows the user to find an optimal value for a single cell, called the target cell, by changing the value of other user-specified cells, called adjustable cells. The target cells and the adjustable cells are directly or indirectly related via formulas. Once the constraints relating to iterations, precision, and convergence are inputted, the program will change the values in the adjustable cells to achieve the desired kind of optimization in the target cell, i.e. maximum, minimum, or target value. In this case a SSR method is computed as the desired minimized target. The residuals are computed as the difference between the measured and theoretical volumetric water contents and summed over all equilibrium pressures (i.e. 0.5 Bar, 1 BAR, 3 BAR, 5 BAR).

#### Solver Setup

All options are selected in a fashion which allows for the best possible precision. The largest values allowed by Solver for iterations and computation time are chosen and the smallest values for precision and tolerance are selected. The default values for all other options are accepted. Table 11 describes the necessary Solver inputs relating to iterations, precision, and convergence.

**Table 11: Excel Solver Configuration**

User Input	Value
Maximum Computation time	10,000 sec
Iterations	10,000 sec
Precision	0.0000001
Tolerance	0.0000001 %
Convergence	0.0000001
Other Constraints	None

A copy of the final spreadsheet used and the interdependent relationships between cells is given in Appendix I. It is important to note that CHEMFLO-2000 specifies input for Van Genuchten's alpha parameter to be 1/cm in the form of:

$$\Theta(h) = \Theta_r + \frac{\Theta_s - \Theta_r}{\left[1 + (-\alpha h)^n\right]^m}$$

while the for Van Genuchten relationship is commonly expressed as:

$$\Theta(\psi) = \Theta_r + \frac{\Theta_s - \Theta_r}{\left[1 + \left(\frac{\alpha}{\psi}\right)^n\right]^m}$$

The bottom equation was also the form used in parameter optimization. This difference between the two forms is accounted for by taking the inverse of the parameter optimization results for input into CHEMFLO. The negative alpha in CHEMFLO's representation of the VG model is accounted for in the alternate VG form for as the difference between matric suction, a positive number, and matric potential, a negative number. It is also important to note the VG parameter  $m$  did not need to be optimized because it is not an independent variable, but instead is calculated from the optimized parameter  $n$ .

### 6.2.2 Parameter Optimization Results

The Parameter Optimization process resulted in the soil-water characteristic function parameters given in Table 12.

**Table 12: Parameter Optimization Results**

Model	Parameter	Physical Interpretation	Suggested limits by CHEMFLO	Parameter Optimization Results		% Difference From Limits
Van Genuchten	$\alpha$ (1/cm)	Empirical constant	$\sim 0.002 < \alpha < \sim 0.15$	Water	0.003901	0.0%
				Leachate	0.003747	0.0%
	n	Empirical constant	$\sim 1.25 < n < 3.0$	Water	1.103105	12.5%
				Leachate	1.129887	10.1%
Brooks Corey	$h_b$ (cm)	Air-entry value	$\sim 1.0 < h_b < \sim 60$	Water	149.1409	85.2%
				Leachate	158.9680	90.4%
	$\lambda$	Empirical constant	$\sim 1.25 < n < 3.0$	Water	0.091387	3.9%
				Leachate	0.109697	3.8%

### **6.3 CHEMFLO – 2000 Methods**

Finite difference techniques are used in this program to solve the partial differential equations for the associated initial and boundary conditions. This method approximates the derivatives in the equations by difference equations in both time and position coordinates. Solutions are then available only at these selected positions and times.

Because the purpose of this research is to discuss the impacts that leachate might have on the SWCC and consequent flow conditions through an unsaturated clay as it pertains to landfill liners, every effort was made to model the soil in terms of realistic landfill conditions where feasible. Decisions about modeling inputs such as initial conditions, boundary conditions, and soil column properties are made to match field conditions. CHEMFLO is used for three modeling scenarios: Fixed Conductivity, Non-Fixed Conductivity, and Normalized Parameters. The goal of each of these scenarios is to highlight leachate's affect on both Van Genuchten and Brooks Corey SWCC on pore fluid interactions and flow characteristics.

#### **Fixed Conductivity**

In this scenario a fixed conductivity function is used in to solve Richards equation for changes in matric potential, hydraulic conductivity, and cumulative flux. The fixed conductivity is the Van Genuchten conductivity function with the same optimized parameters from the VG SWCC. The scenario is run once using the VG and BC soil-water characteristic curves. An advantage to this scenario is the affects of leachate on the VG and BC SWCCs can be quantified, and because the SWCC parameters are the only variables the difference in how VG and BC quantify these affects can also be compared.

#### **Non-Fixed Conductivity**

Research has indicated that there is a definite relationship between matric suction and permeability. Even through permeability was not measured as part of this research; a relative conclusion can be made about the affects of Leachate on permeability by changing the conductivity parameters relative to how they have changed in the SWCC. In this scenario the conductivity function is not fixed. Instead of using VG parameters

from the water SWCC for all analysis, the VG parameters for leachate are used when modeling leachate.

The rationale for being able to compare the effects on permeability without calibrating permeability parameters comes from a general relationship between the SWCC and permeability function described by Leong and Rhardjo as:

$$k_r = \Theta^P$$

If  $P$  is known for a particular soil, then permeability as a function of matric potential can be derived. In this case,  $P$  is not known, but because the *change* in permeability is of interest and the soil is the same, then difference in permeability is a function of changes in the SWCC only.

### **Normalized Parameters**

Part of the variation in the SWCC curves, and in turn matric potential, hydraulic conductivity, and cumulative flux, is due to experimental variation—in this case compaction conditions. The dissimilarity of the water and leachate soil water characteristic curves is the result of a difference in Van Genuchten parameters and saturated and residual water contents—these are the inputs to Richards' equation and also CHEMFLO. While the difference in VG parameters are the result of varying equilibrium water contents, the saturated and residual volumetric water contents are derived from geotechnical phase relationships which are dependent on the average compacted dry density for the soil samples. Under ideal laboratory conditions, compacted dry density would be the same for samples permeated with leachate and water; however, slight differences in compaction conditions will result in different saturated and residual water contents.

In order to quantify the effects of variation in compaction conditions, this scenario normalizes  $\Theta_s$ ,  $\Theta_r$ , and dry density by taking the average of these variables between leachate and water for input into CHEMFLO.

#### **6.4 Inputs to CHEMFLO**

Inputs to CHEMFLO for soil conditions and SWCC and conductivity parameters are summarized in Table 13 and

Table 14 for each modeling scenario for the Van Genuchten and Brooks Corey models, respectively.

Table 13: Soil System Input to CHEMFLO for Van Genuchten SWCC

	CHEMFLO Input	Water	Leachate	How data is utilized	Data Source
Fixed Conductivity	<b>Soil Properties</b>				
	$\Theta_s$	0.4576	0.4555	SWCC and Permeability Functions	Average of all $\Theta_s$ as calculated from each compacted sample
	$\Theta_r$	0.0752	0.0755	SWCC and Permeability Functions	Average of all $\Theta_r$ as calculated from each compacted sample
	Dry Density	1.4493 24	1.45473 3	SWCC and Permeability Functions	Average of all Dry Density for each sample
	<b>Van Genuchten SWCC Properties</b>				
	$\alpha$ (1/cm)	0.0039 01	0.00374 7	VG SWCC Function	VG SWCC Parameter optimization
	n	1.1031 05	1.12988 7	VG SWCC Function	VG SWCC Parameter Optimization
	<b>Conductivity Function Properties</b>				
	$K_s$ (cm/hour)	3.64E-4	3.64E-4	VG Permeability Function	Previous engineering reports on the Chicago Grade “brown matrix” soil.
	$\alpha$ (1/cm)	0.0039 01	0.00390 1	VG Permeability Function	VG SWCC Parameter optimization - Water
Not-Fixed Conductivity	n	1.1031 05	1.10310 5	VG Permeability Function	VG SWCC Parameter Optimization - Water
	<b>Soil Properties</b>				
	$\Theta_s$	0.4576	0.4555	SWCC and Permeability Functions	Average of all $\Theta_s$ as calculated from each compacted sample
	$\Theta_r$	0.0752	0.0755	SWCC and Permeability Functions	Average of all $\Theta_r$ as calculated from each compacted sample
	Dry Density	1.4493 24	1.45473 3	SWCC and Permeability Functions	Average of all Dry Density for each sample
	<b>Van Genuchten SWCC Properties</b>				
	$\alpha$ (1/cm)	0.0039 01	0.00374 7	VG SWCC Function	SWCC Parameter optimization
	n	1.1031 05	1.12988 7	VG SWCC Function	SWCC Parameter Optimization
	<b>Conductivity Function Properties</b>				
	$K_s$ (cm/hour)	3.64E-4	3.64E-4	VG Permeability Function	Previous engineering reports on the Chicago Grade “brown matrix” soil.
	$\alpha$ (1/cm)	0.0039 01	0.00374 7	VG Permeability Function	VG SWCC Parameter optimization – Leachate and Water
	n	1.1031 05	1.12988 7	VG Permeability Function	VG SWCC Parameter Optimization – Leachate and

	CHEMF LO Input	Water	Leachate	How data is utilized	Data Source
					Water
Normalized Parameters ( Fixed Cond.)	Soil Properties				
	$\Theta_s$	0.45655	0.45655	SWCC and Permeability Functions	Average of all $\Theta_s$ for Leachate and Water
	$\Theta_r$	0.07537	0.07537	SWCC and Permeability Functions	Average of all $\Theta_r$ for Leachate and Water
	Dry Density	1.452029	1.452029	SWCC and Permeability Functions	Average of all Leachate and Water Samples
	Van Genuchten SWCC Properties				
	$\alpha$ (1/cm)	0.003901	0.003747	VG SWCC Function	SWCC Parameter optimization
	n	1.103105	1.129887	VG SWCC Function	SWCC Parameter Optimization
	Conductivity Function Properties				
	Ks (cm/hour)	3.64E-4	3.64E-4	VG Permeability Function	Previous engineering reports on the Chicago Grade "brown matrix" soil.
	$\alpha$ (1/cm)	0.003901	0.003747	VG Permeability Function	VG SWCC Parameter optimization - Water
	n	1.103105	1.129887	VG Permeability Function	VG SWCC Parameter Optimization - Water

Table 14: Soil System Input to CHEMFLO for Brooks Corey SWCC

	CHEMFLO Input	Water	Leachate	How data is utilized	Data Source
Fixed Conductivity	Soil Properties				
	$\Theta_s$	0.4576	0.4555	SWCC and Permeability Functions	Average of all $\Theta_s$ as calculated from each compacted sample
	$\Theta_r$	0.0752	0.0755	SWCC and Permeability Functions	Average of all $\Theta_r$ as calculated from each compacted sample
	Dry Density	1.4493 24	1.45473 3	SWCC and Permeability Functions	Average of all Dry Density for each sample
	Brooks Corey SWCC Properties				
	$h_b$ (cm)	149.14 0866	158.968 044	BC SWCC Function	BC SWCC Parameter optimization
	$\lambda$	0.0913 87	0.10969 7	BC SWCC Function	BC SWCC Parameter Optimization
	Conductivity Function Properties				
	$K_s$ (cm/hour)	3.64E-4	3.64E-4	VG Permeability Function	Previous engineering reports on the Chicago Grade "brown matrix" soil.
	$\alpha$ (1/cm)	0.0039 01	0.00390 1	VG Permeability Function	VG SWCC Parameter optimization - Water
Not-Fixed Conductivity	$n$	1.1031 05	1.10310 5	VG Permeability Function	VG SWCC Parameter Optimization - Water
	Soil Properties				
	$\Theta_s$	0.4576	0.4555	SWCC and Permeability Functions	Average of all $\Theta_s$ as calculated from each compacted sample
	$\Theta_r$	0.0752	0.0755	SWCC and Permeability Functions	Average of all $\Theta_r$ as calculated from each compacted sample
	Dry Density	1.4493 24	1.45473 3	SWCC and Permeability Functions	Average of all Dry Density for each sample
	Brooks Corey SWCC Properties				
	$h_b$ (cm)	149.14 0866	158.968 044	BC SWCC Function	BC SWCC Parameter optimization
	$\lambda$	0.0913 87	0.10969 7	BC SWCC Function	BC SWCC Parameter Optimization
	Conductivity Function Properties				
	$K_s$ (cm/hour)	3.64E-4	3.64E-4	VG Permeability Function	Previous engineering reports on the Chicago Grade "brown matrix" soil.
	$\alpha$ (1/cm)	0.0039 01	0.00374 7	VG Permeability Function	VG SWCC Parameter optimization – Leachate and Water
	$n$	1.1031 05	1.12988 7	VG Permeability Function	VG SWCC Parameter Optimization – Leachate and

	CHEMFL O Input	Water	Leachate	How data is utilized	Data Source
					Water
Normalized Parameters (Fixed Cond.)	<b>Soil Properties</b>				
	$\Theta_s$	0.4565 5	0.45655	SWCC and Permeability Functions	Average of all $\Theta_s$ for Leachate and Water
	$\Theta_r$	0.0753 7	0.07537	SWCC and Permeability Functions	Average of all $\Theta_r$ for Leachate and Water
	Dry Density	1.4520 29	1.45202 9	SWCC and Permeability Functions	Average of all Leachate and Water Samples
	<b>Brooks Corey SWCC Properties</b>				
	$h_b$ (cm)	149.14 0866	158.968 044	BC SWCC Function	BC SWCC Parameter optimization
	$\lambda$	0.0913 87	0.10969 7	BC SWCC Function	BC SWCC Parameter Optimization
	<b>Conductivity Function Properties</b>				
	$K_s$ (cm/hour)	3.64E-4	3.64E-4	VG Permeability Function	Previous engineering reports on the Chicago Grade “brown matrix” soil.
	$\alpha$ (1/cm)	0.0039 01	0.00374 7	VG Permeability Function	VG SWCC Parameter optimization - Water
	$n$	1.1031 05	1.12988 7	VG Permeability Function	VG SWCC Parameter Optimization - Water

#### 6.4.1 Initial Conditions

The software can simulate water movement in soil columns with uniform or non-uniform initial conditions. It can also simulate water movement in semi-infinite soils with uniform initial conditions. In this study a dimension of length of “L” extending in the positive y direction was selected to simulate a vertical, finite soil column. Therefore:

$$h(y,0) = h_{\text{initial}}(y) \text{ where } 0 < y < L$$

Liner thickness is usually specified between five to twelve feet of clay with two feet as the minimum by EPA guidance (Clay Liners for Waste Management Facilities, pg 258). The minimum thickness of two feet (60.96 cm) was used in the computer analysis. This lower bound value was chosen in order to minimize processing time while at the same time simulating an appropriate thickness.

The program will compute initial total potential by imputing initial volumetric water content -- computed from 2% above optimum water content as measured in the laboratory. 25.5% plus 2% = 27.5%. Average volumetric water content is computed by taking the average of computed volumetric water contents for leachate and water at 27.5%-- they differ slightly because of the effect of dry density in the computation of volumetric water content—which is 0.399.

#### 6.4.2 Initial Boundary conditions

Three boundary conditions can be applied to the soil at  $y = 0$ :

1. Constant potential of  $h_o$
2. Constant flux density of  $q_o$
3. Mixed type where a constant flux density can be specified until a user-specified critical matric potential is reached at which time  $h_c = h_o$
4. Falling Head where the user specifies the duration and depth of ponding.

For this study a falling head scenario is specified for a given ponding time. Therefore:

$$h(0,0) = h_o \text{ where } h_o > 0$$

$$\frac{dh}{dt} = -q(0,t) \text{ for } 0 < t < t_p$$

$$\text{and where } q_o = -K(H) \left( \frac{\partial h}{\partial y} - \sin(A) \right)_{y=0}$$

$$q(0,t) = 0 \text{ for } t \geq t_p$$

$$\text{and } (h,0) > 0 \text{ for } t < t_p$$

This would simulate ponding and provide a driving force that would lead to leakage in a clay liner. The initial ponding depth is arbitrarily set equal to twenty cm. Setting the initial ponding depth higher than this created large mass balance errors. An explanation for this is the model sees this as an instantaneous change in head and accompanying this head change will be an instantaneous change in fluid stored within an element. Since this

increase (or decrease) in storage was not accounted for in the previous time step the model treats it as a discrepancy in the total mass balance

#### 6.4.3 Final Boundary Conditions

For a finite soil system, one of the following boundary conditions can be imposed at  $y = L$ .

1. Constant matric potential of  $h_L$
2. Constant flux density of  $q_L$
3. Free drainage

For this study free drainage is assumed where

$$\left. \frac{\partial h}{\partial x} \right|_{x=L} = 0$$

This condition is appropriate as the cumulative flux through the soil column is of primary concern.

#### 6.4.4 Mesh Size and Convergence

A collection of points at which solutions exist are called mesh points. The distance between mesh points is commonly called the mesh size. CHEMFLO uses a uniform mesh size in position within each layer of soil. However, mesh sizes may differ between layers. The solution obtained by this technique depends to some extent upon the mesh sizes used. In general, smaller mesh sizes produce more accurate results.

Richards's equation is highly nonlinear and solving it requires an iterative approach. Specifying the convergence criteria is a critical component to a numerical modeling and specifying incorrect or inappropriate mesh sizes will result in unreliable solutions. A numerical solution to Richards Equation requires input regarding the:

1. time mesh
2. distance mesh
3. convergence

CHEMFLO calculates maximum absolute errors relative mass balance errors. The absolute mass balance error is calculated as the difference in net amount of water entering the soil and the increase in the amount of water stored in the soil (Nofziger and Wu 2003). This difference divided by the amount of water entering the soil represents the relative mass balance error. CHEMFLO provides the user with the ability to set the limits on the maximum absolute error and relative error. If during the course of a computation the program encounters errors greater than the user-defined limits, the program will warn the user accordingly.

Smaller mesh sizes are intuitively desirable because they are associated with a more exact answer than larger meshes. A mesh size/convergence study was conducted to determine the convergence criteria and the range of mesh sizes which could be used without invoking convergence or mass balance errors. In order to do this, the optimized parameters and boundary conditions were inputted similar to how they would be in a modeling scenario and the mesh distances, time meshes, and convergence criteria were varied in order to produce an array of acceptable mesh sizes. This mesh size/convergence study was completed for both water and leachate samples, and both sample resulted in the same mesh size and convergence criteria limits. The results are summarized in Table 15 and complete results are presented in Appendix III.

**Table 15: Results of Mesh/Convergence**

Criteria	Minimum Value
Distance Mesh (cm)	0.000001
Time Mesh (hours)	0.1
Convergence	0.01
Maximum Relative Error (cm)	0.125
Maximum Absolute Error (%)	5.0
Critical Cumulative Flux Error (cm)	2.5

## 7 Results

The results of this study can be divided into three categories:

1. Statistical Significance of Experimental Data.
2. Effects of Leachate on the Soil Water Characteristic Curve using the Van Genuchten and Brooks Corey Models. This result is discussed in terms of how parameters have affected common properties of the soil water characteristic curve including air-entry suction head.
3. Quantitative effects of Leachate. Comparison of unsaturated fluid of movement of leachate verses water in terms of matric potential and cumulative flux through a soil column using the Van Genuchten and Brooks Corey Model.

Clearly these results have not been achieved independently; rather the affects of different permeates on the soil water characteristic curve will carry through to the quantitative comparison of unsaturated fluid movement. However, it is important to highlight these results separately. Basic results and output from CHEMFLO are presented in this section and then are discussed in Section 8.

### 7.1 Result 1 – Statistical Significance

There is an apparent decline in measured equilibrium water contents for samples permeated with leachate. In order to determine if this decline is significant or the result of experimental variability, a two-tailed T-test is conducted on the difference in means for the final measured gravimetric water contents. The null hypothesis is no difference exists between the mean water contents for leachate and water at a given equilibrium moisture content. The outcome is that experimental results are significant at the 2.5% level, meaning the probability of measuring a difference similar to what was measured in this study if leachate had no affect on equilibrium water contents is less than 2.5%. The T-test results are summarized in Table 17. Table 16: Experimental Data Results

Matric Suction	Water			Leachate			Difference in Means
	Final Measured Gravimetric Water Content	Variance	Mean	Final Measured Gravimetric Water Content	Variance	Mean	
0.5-BAR	0.2878	1.055E-06	0.2886	0.2761	7.305E-06	0.2792	0.0094
	0.2883			0.2807			
	0.2898			0.2808			
1-BAR	0.2692	1.049E-05	0.2729	0.2692	4.450E-07	0.2699	0.0031
	0.2744			0.2700			
	0.2752			0.2705			
3-BAR	0.2526	1.274E-08	0.2525	0.2452	1.030E-05	0.2423	0.0102
	0.2524			0.2388			
	0.2525			0.2429			
5-BAR	0.2442	3.868E-06	0.2436	<sup>1</sup> 0.2293	1.934E-06	0.2284	0.0152
	0.2414			0.2268			
	0.2452			0.2291			

**Table 17: T-Test Results**

Statistical Characteristic	Value
std. deviation	0.004977119
average	0.009465
t-statistic	4.2703
Degrees of Freedom	3
standard error of difference	0.002
<b>two-tailed P value</b>	<b>0.0238</b>

## 7.2 Result 2 – Soil Water Characteristic Curve

The following figures present the SWCC for leachate and water based on the results of the parameter optimization process. Figure 6 through Figure 9 show the VG and BC SWCCs for water and leachate, while Figure 10 shows all curves on one graph. Both models showed a relatively good fit of the data and a close fit to one another.

<sup>1</sup> This sample reached equilibrium pressure at 4.92 Bars, instead of 5.0 Bars.

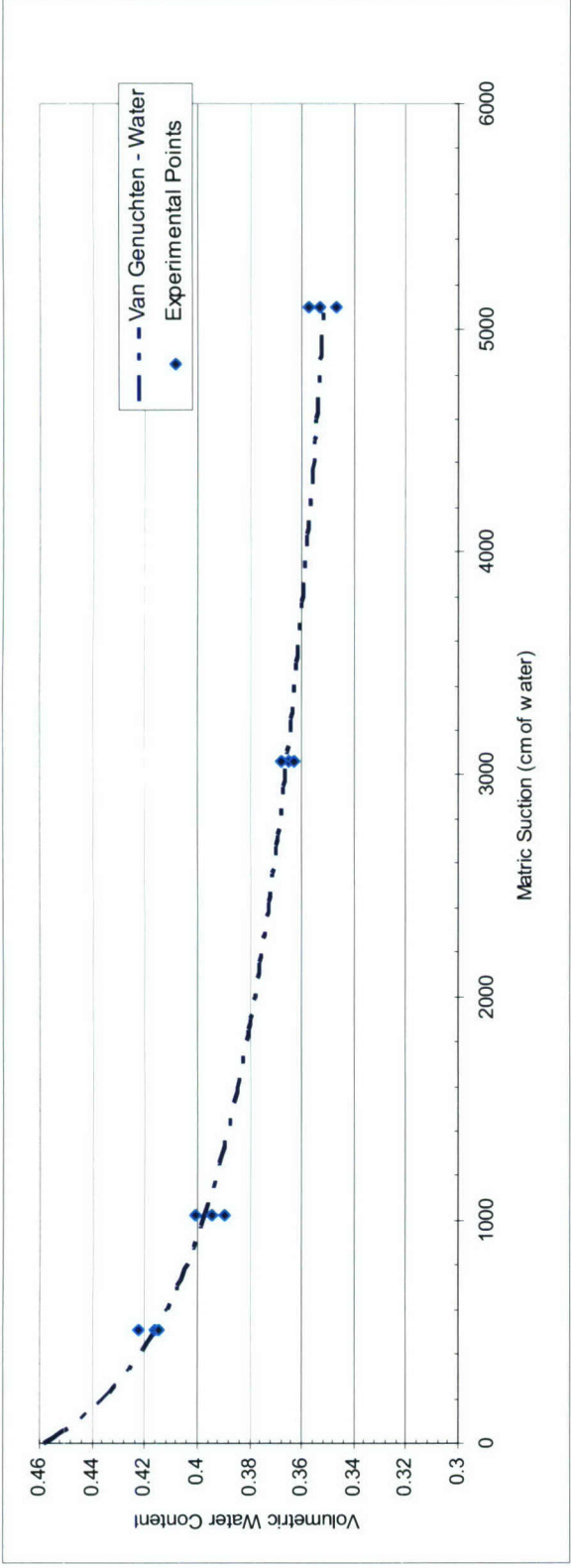


Figure 6: Van Genuchten SWCC for Water

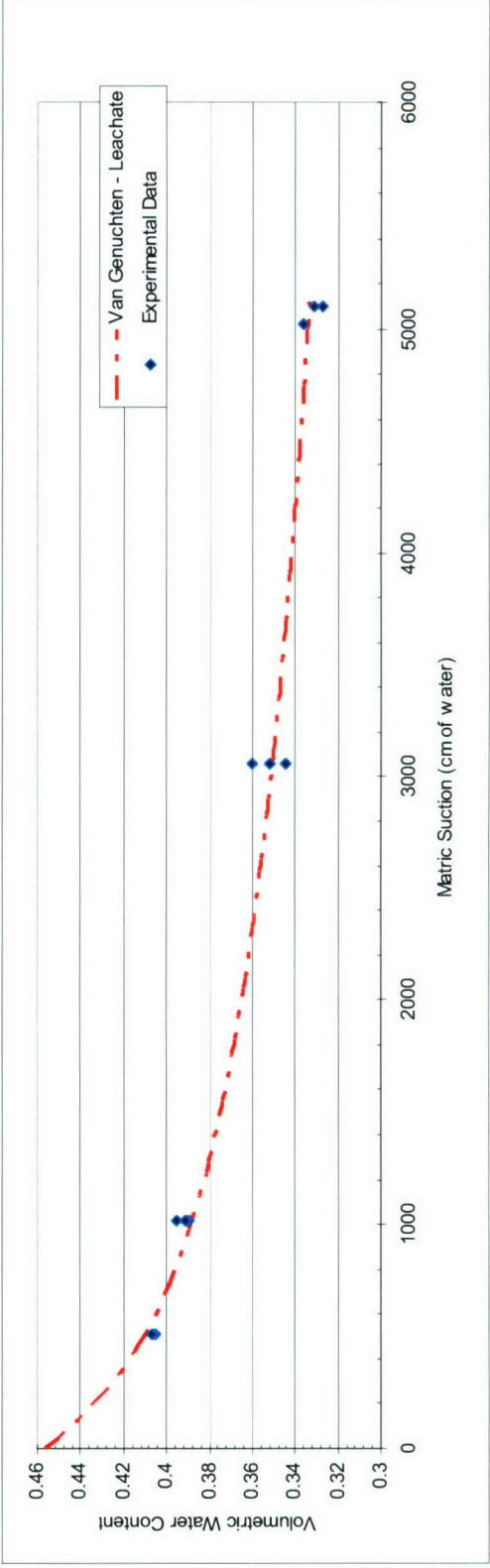


Figure 7: Van Genuchten SWCC for Leachate



**Figure 8: Brooks Corey SWCC for Water**

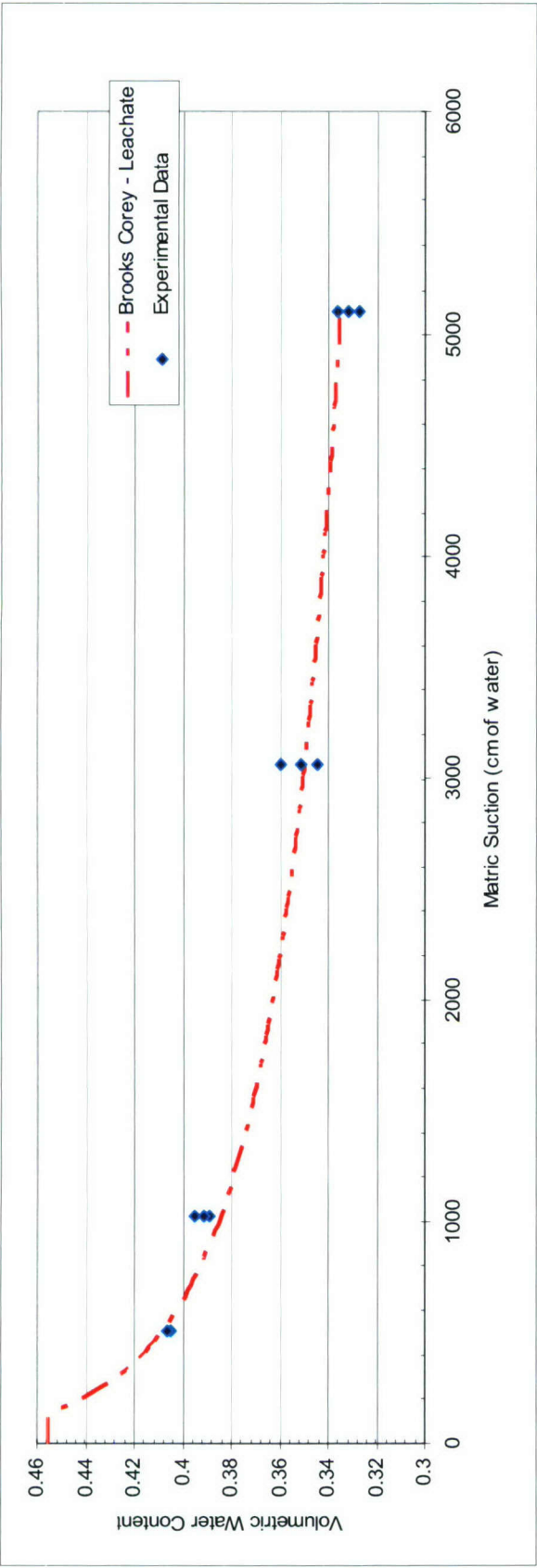


Figure 9: Brooks Corey SWCC for Leachate

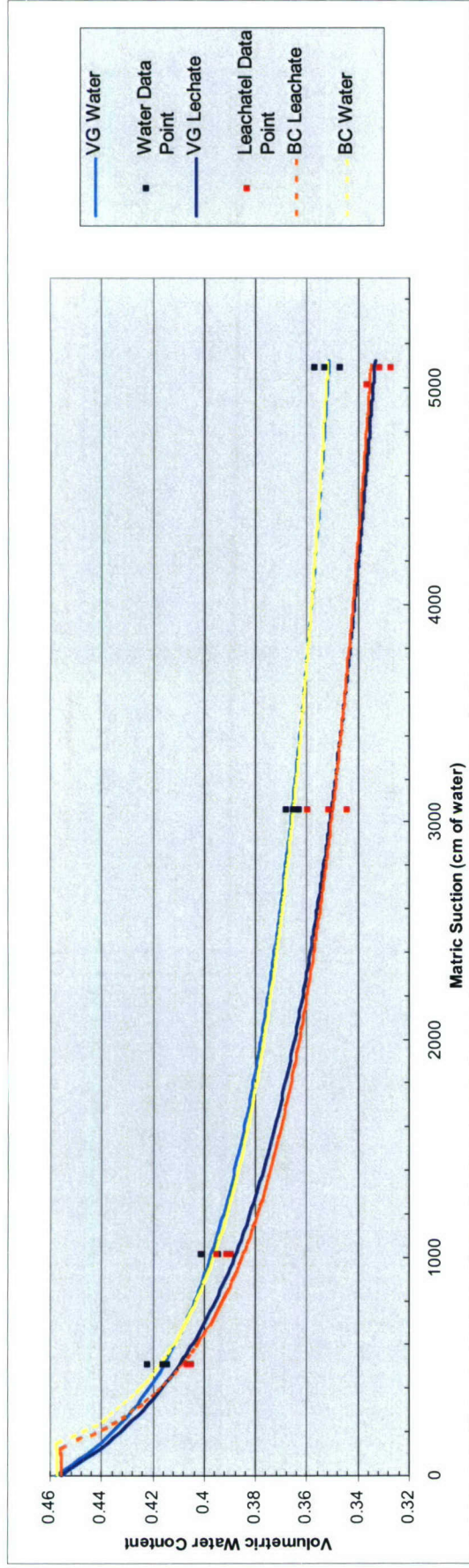


Figure 10: Comparison of Brooks Corey and Van Genuchten SWCC for Leachate and Water

### 7.3 Result 3 – Quantitative Comparison of Effects of Leachate

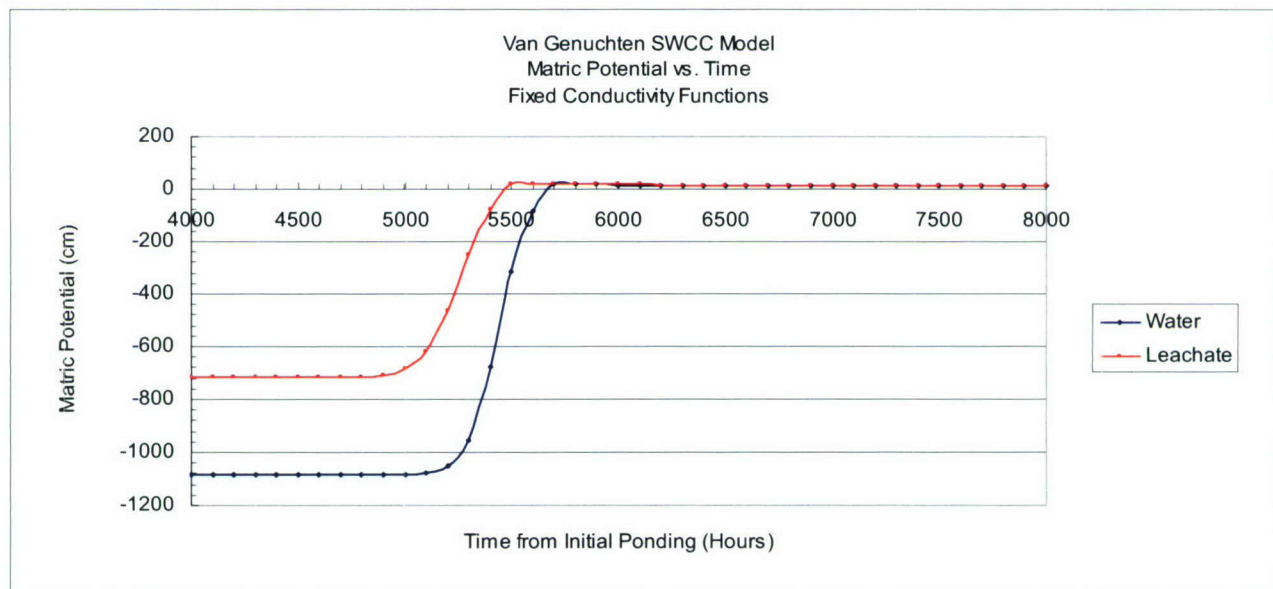
This section will quantify the affects changes in the SWCC have on common soil flow characteristics through an unsaturated soil column using CHEMFLO-2000. As described in Section 6, a soil column is modeled in CHEMFLO using the results of parameter optimization for both the Van Genuchten and Brooks Corey models. In all CHEMFLO scenarios, the results are relative to a single point at the bottom of the 61.0 cm soil column.

Three scenarios have been run in CHEMFLO:

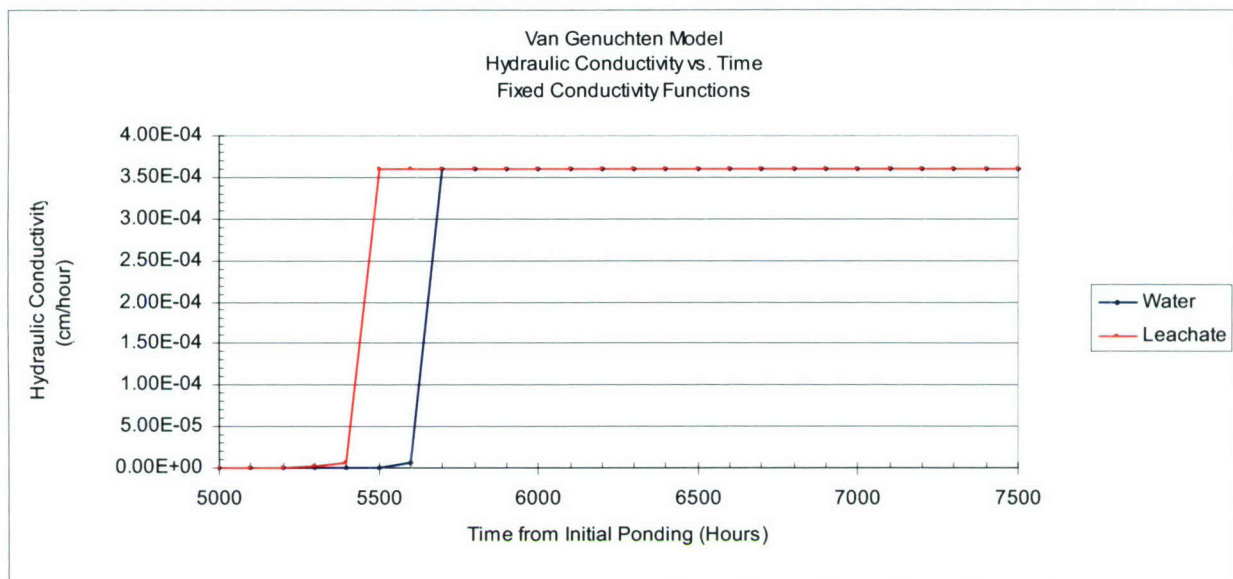
1. Fixed Conductivity
2. Independent Conductivity
3. Normalized  $\Theta_s$ ,  $\Theta_r$ , and dry densities parameters with Fixed Conductivity.

#### 7.3.1 Modeling Scenario #1: Fixed Conductivity

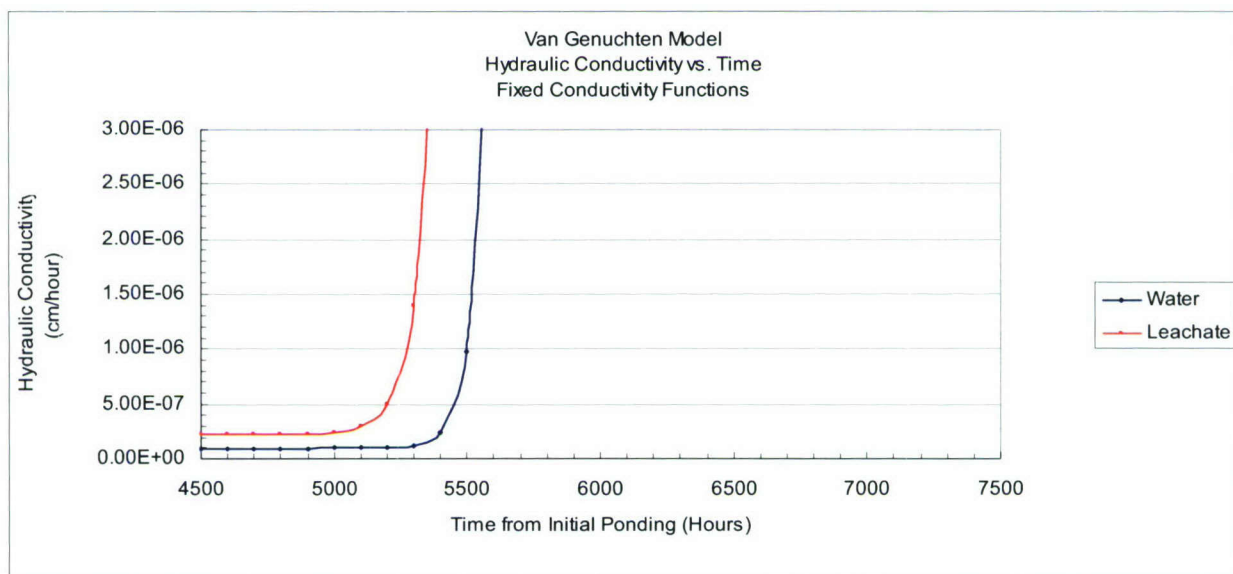
##### Van Genuchten



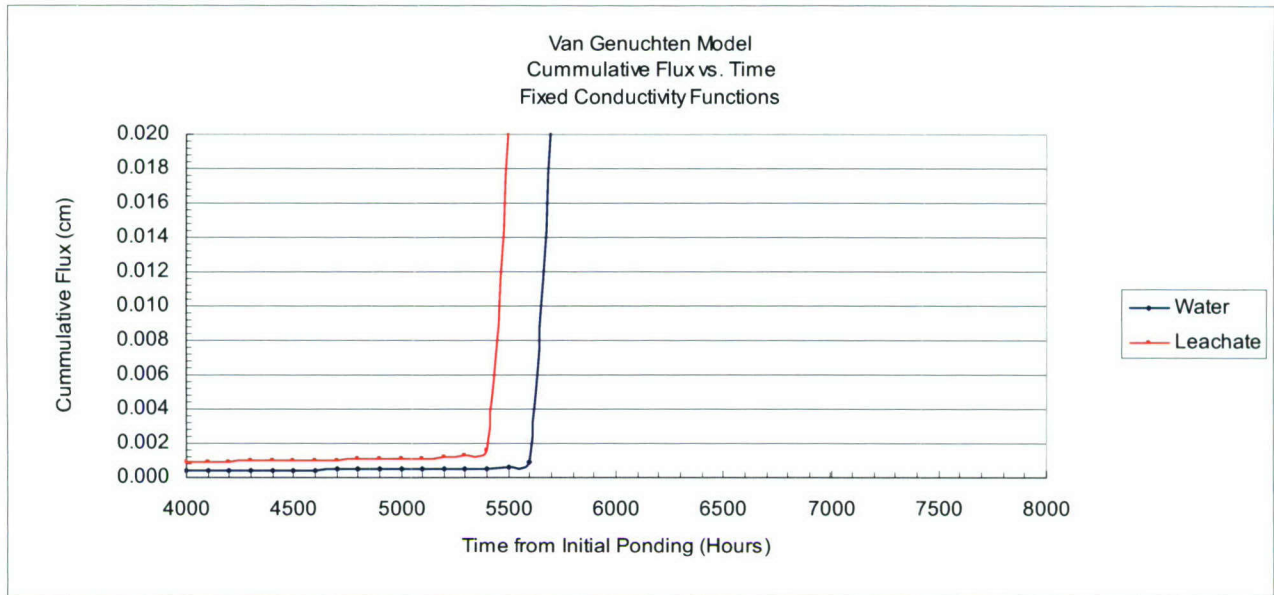
**Figure 11: Comparison of Van Genuchten Predicted Matric Potentials using Fixed Permeability Functions**



**Figure 12: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Fixed Permeability Functions**

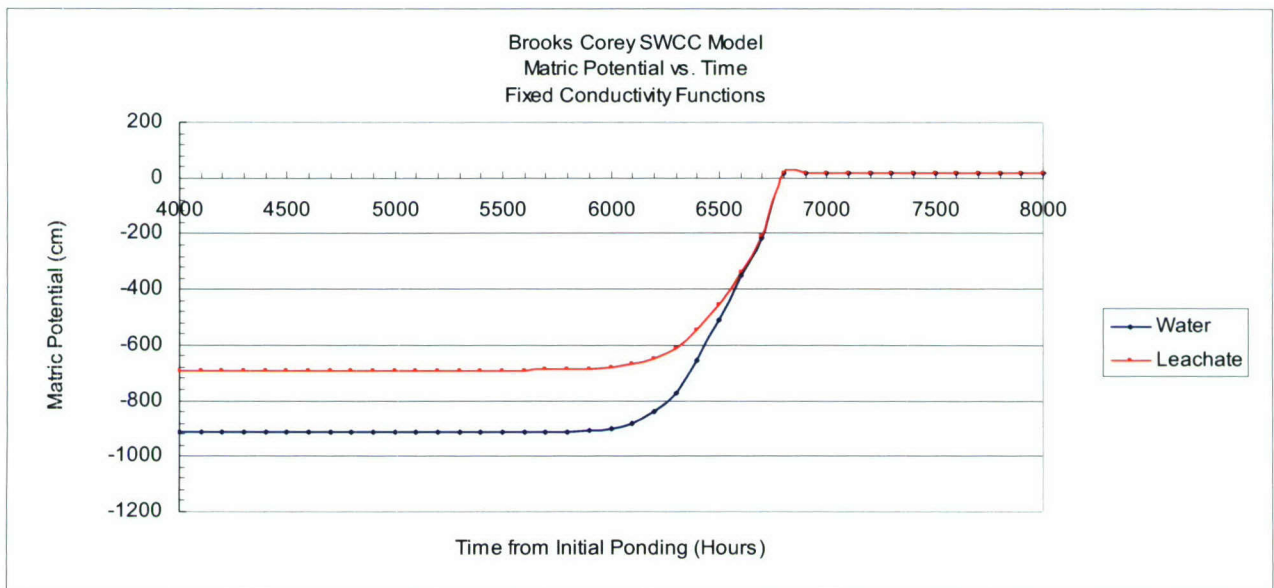


**Figure 13: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Fixed Permeability Functions (Before Saturation)**

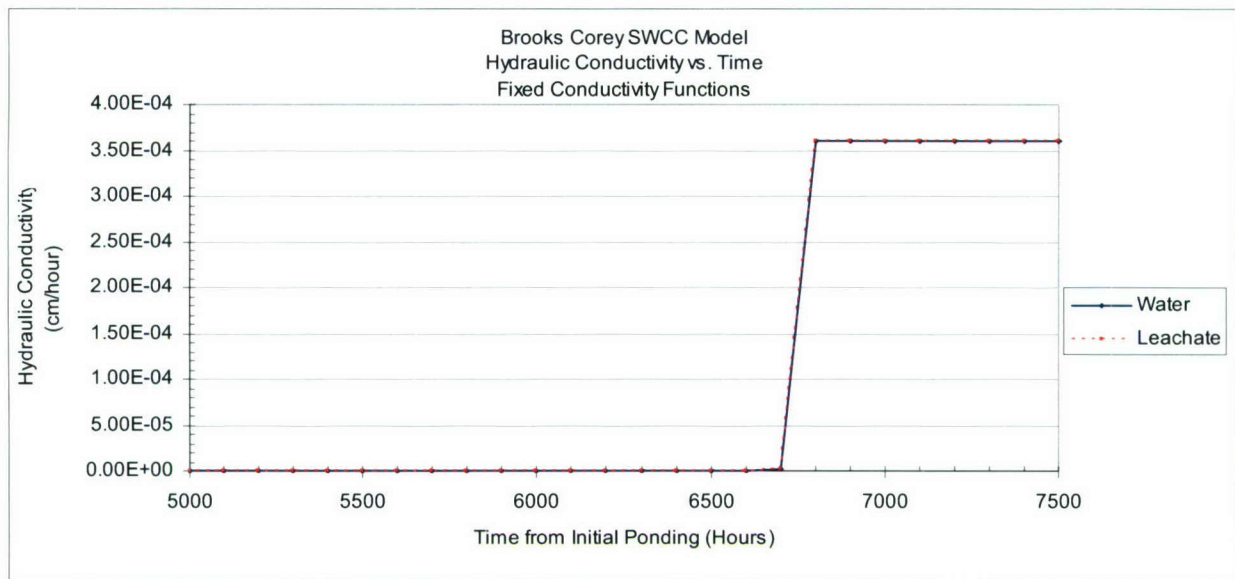


**Figure 14: Comparison of Van Genuchten Predicted Cumulative Flux using Fixed Permeability Functions.**

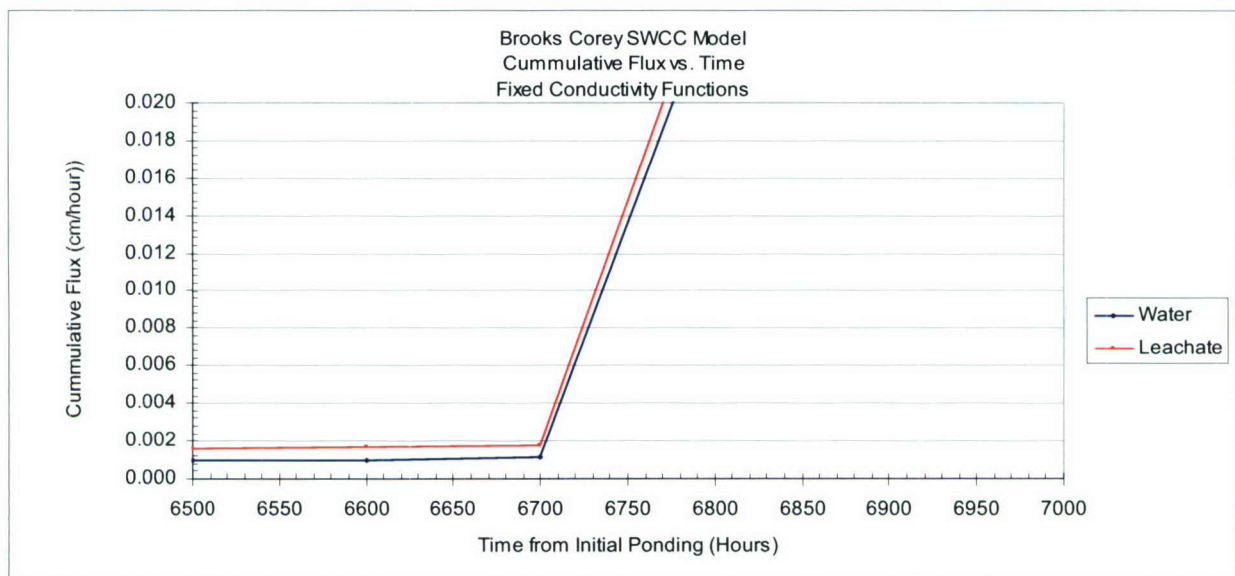
### Brooks Corey



**Figure 15: Comparison of Brooks Corey Predicted Matric Potential using Fixed Permeability Functions.**



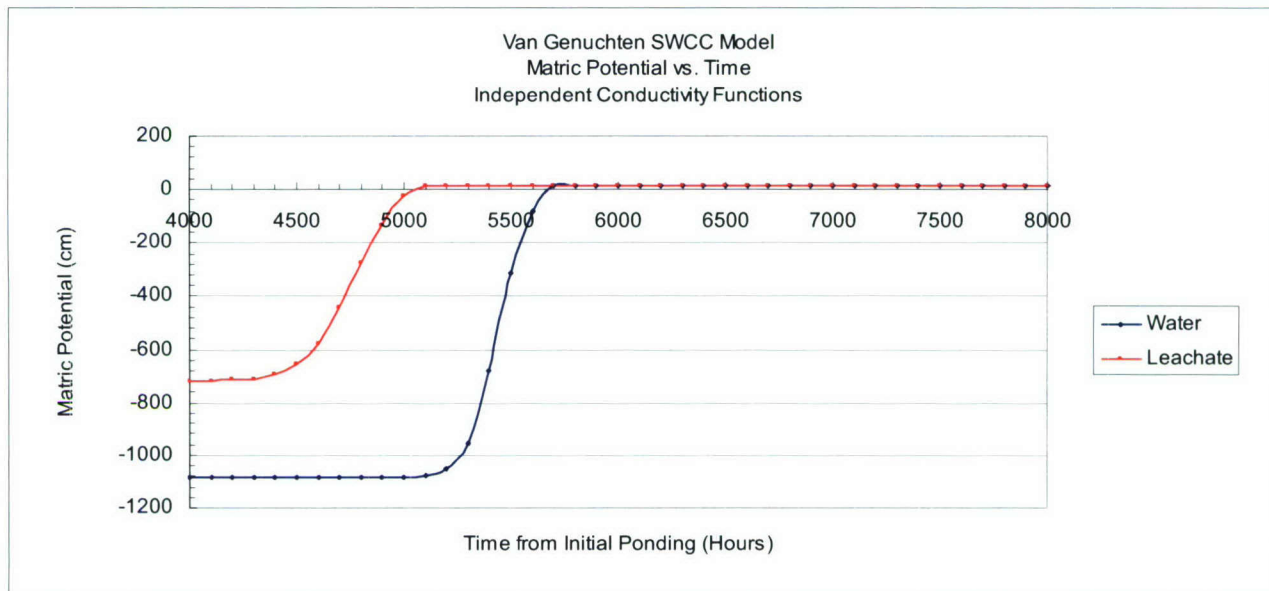
**Figure 16: Comparison of Brooks Corey Predicted Hydraulic Conductivity using Fixed Permeability Functions.**



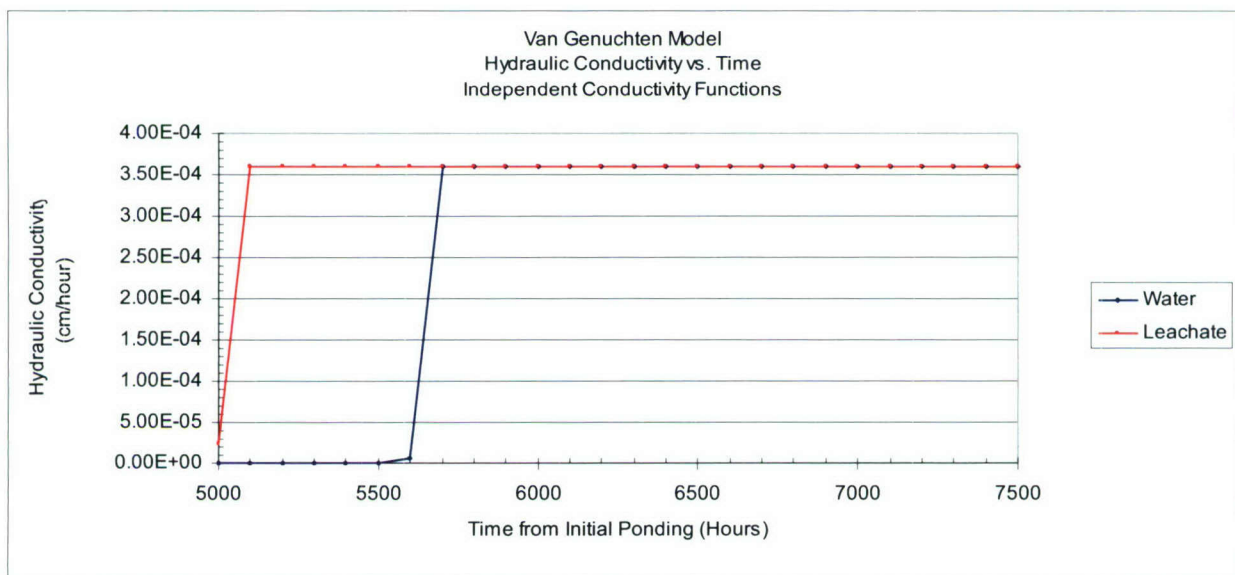
**Figure 17: Comparison of Brooks Corey Predicted Cumulative Flux using Fixed Permeability Functions.**

### 7.3.2 Modeling Scenario #2: Independent Conductivity

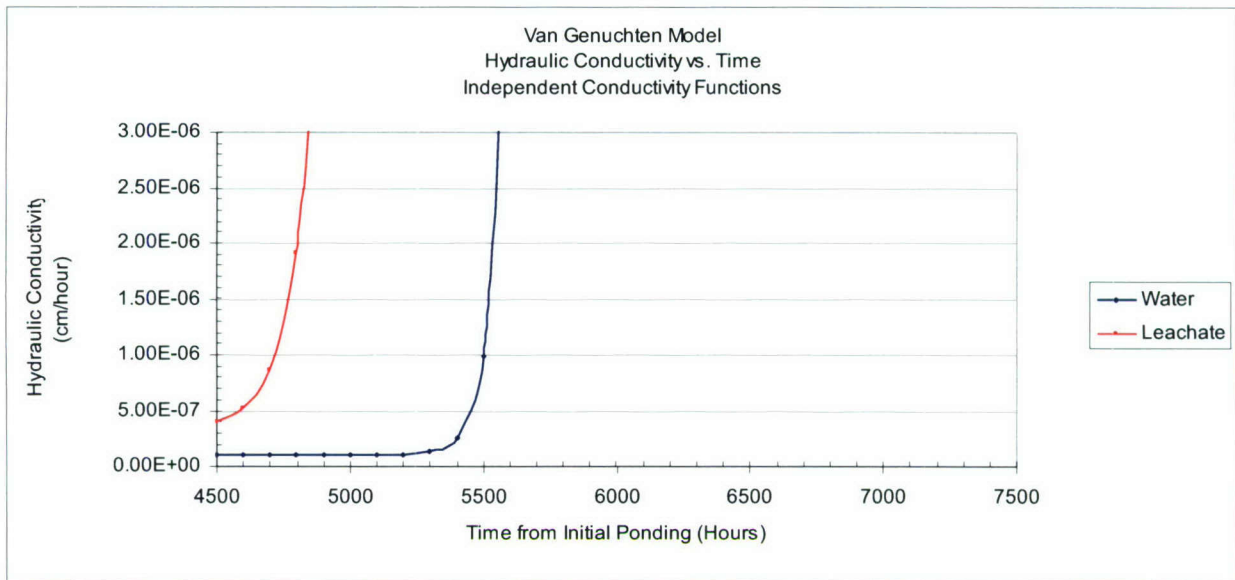
#### Van Genuchten



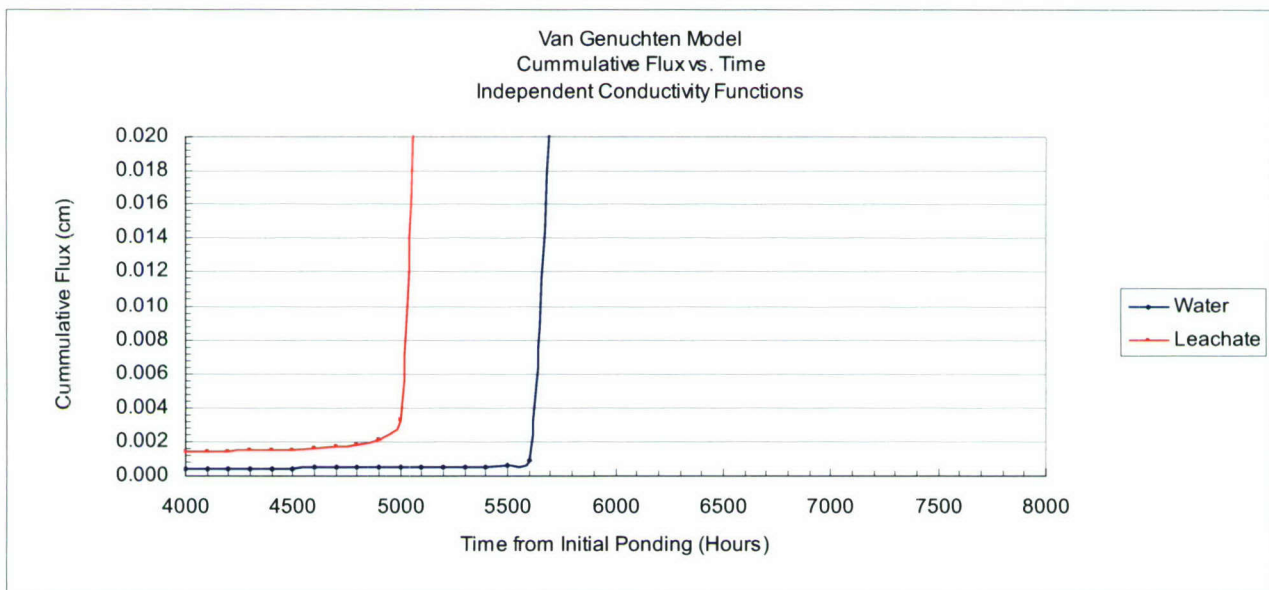
**Figure 18: Comparison of Van Genuchten Predicted Matric Potential using Independent Permeability Functions**



**Figure 19: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Independent Permeability Functions**

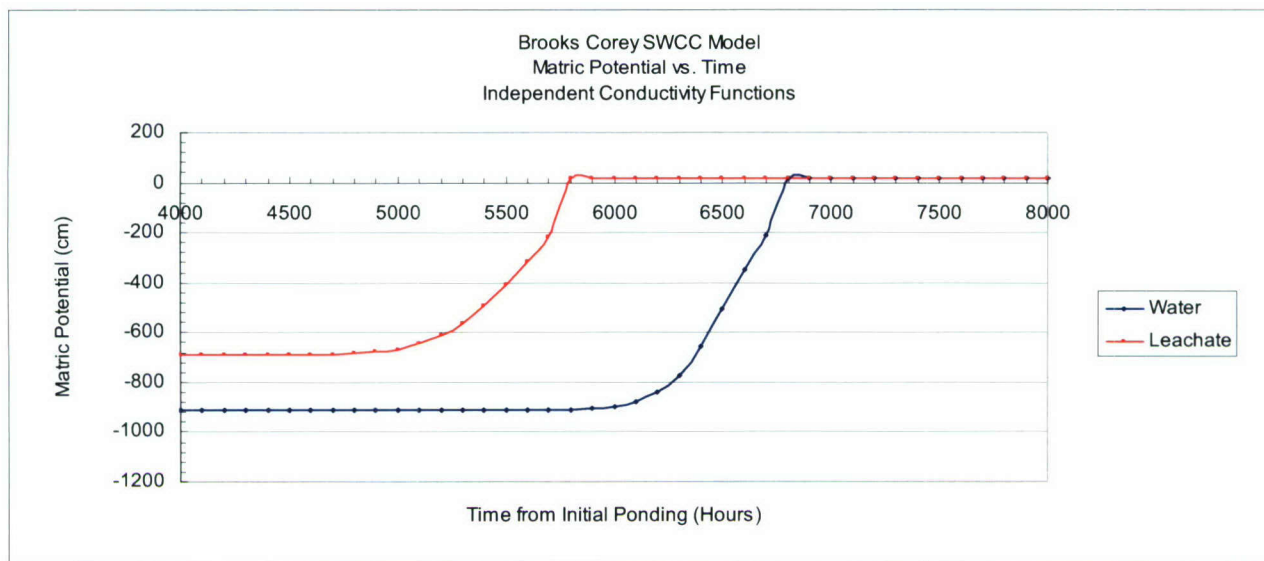


**Figure 20: Comparison of Van Genuchten Predicted Hydraulic Conductivity using Independent Permeability Functions (Before Saturation)**

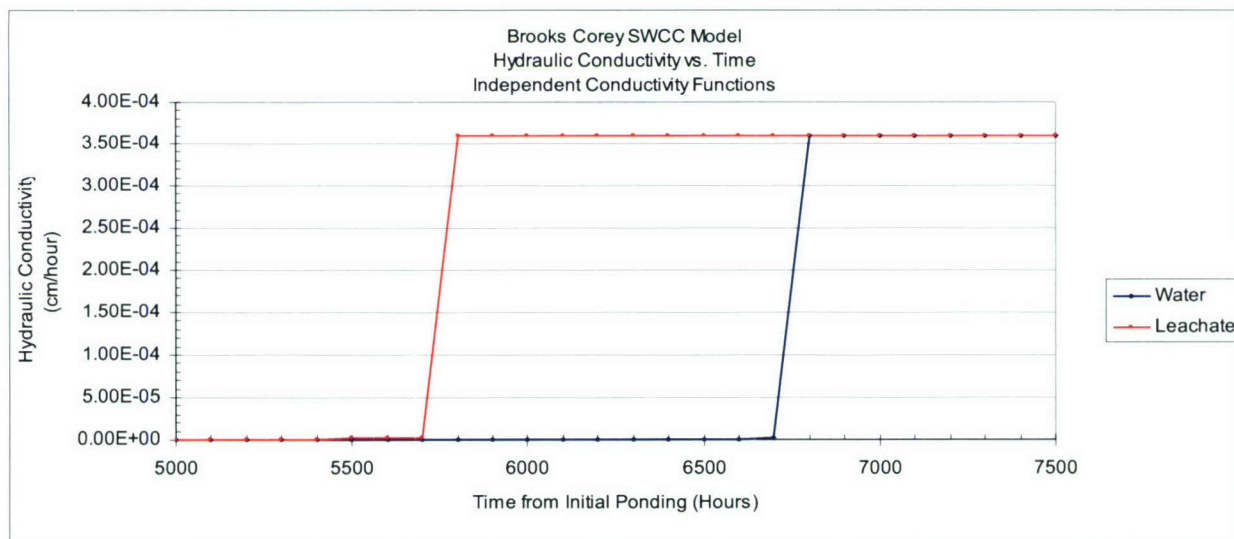


**Figure 21: Comparison of Van Genuchten Predicted Cumulative Flux using Independent Permeability Functions**

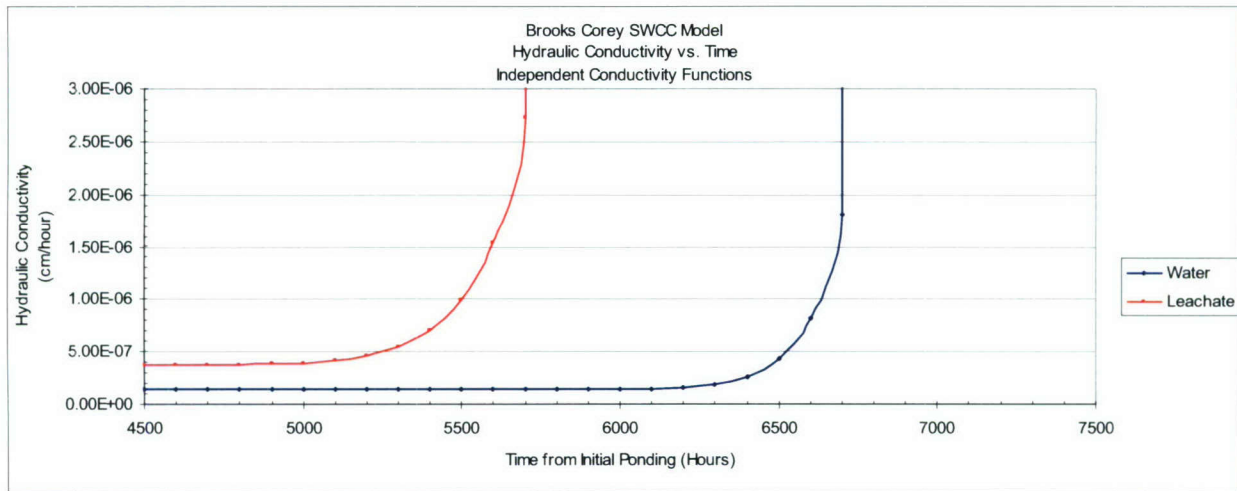
## Brooks- Corey



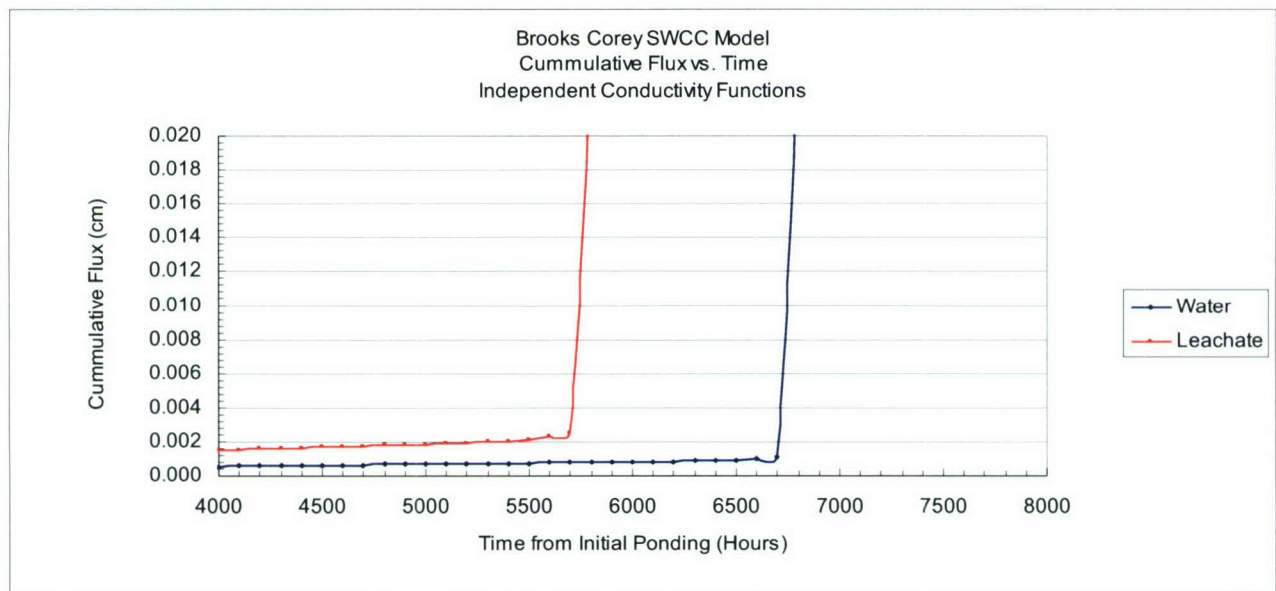
**Figure 22: Comparison of Brooks Corey Predicted Matric Potential using Independent Permeability Functions**



**Figure 23: Comparison of Brooks Corey Predicted Hydraulic Conductivity using Independent Permeability Functions**



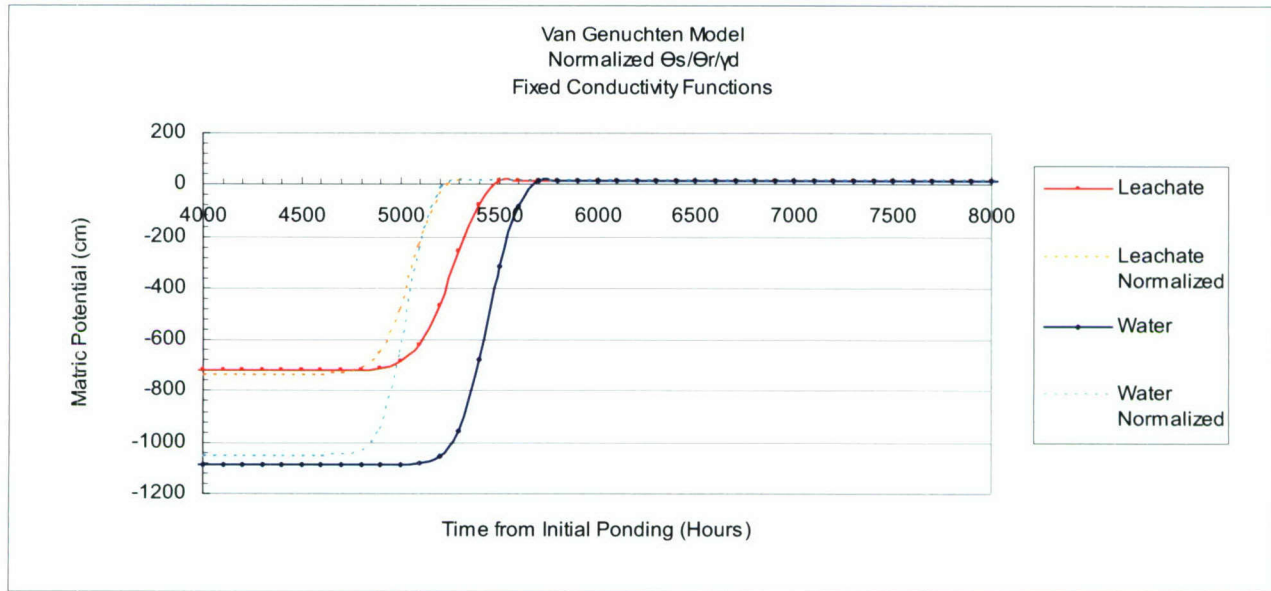
**Figure 24: Comparison of Brooks Corey Predicted Hydraulic Conductivity using Independent Permeability Functions (Before Saturation)**



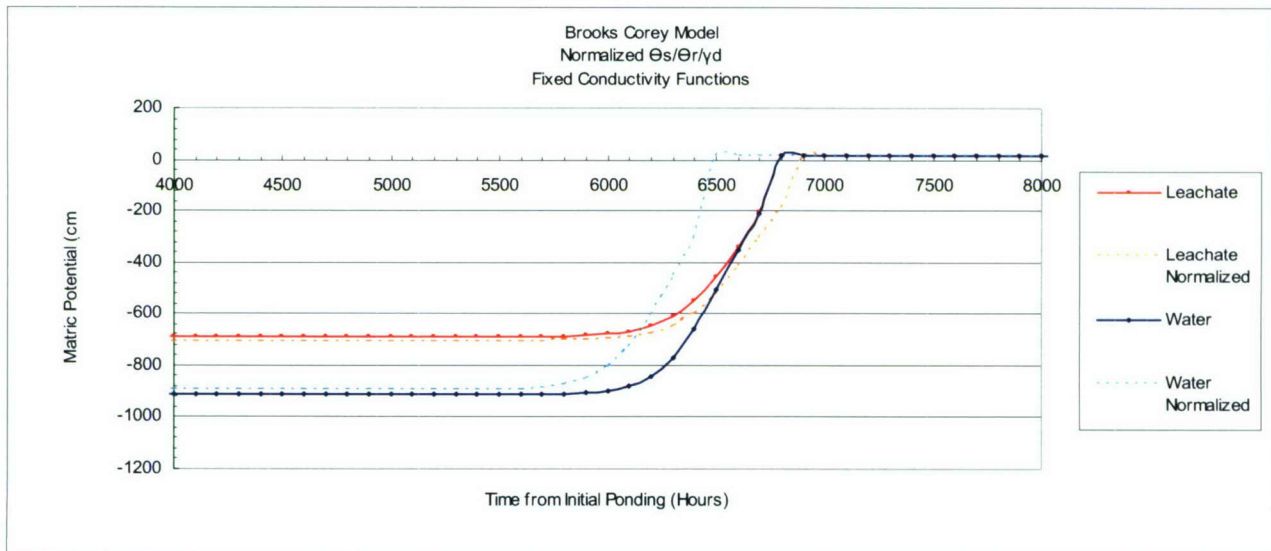
**Figure 25: Comparison of Brooks Corey Predicted Cumulative Flux using Independent Permeability Functions**

### 7.3.3 Modeling Scenario #3: Normalized $\Theta_s$ , $\Theta_r$ , and Dry Densities

#### Van Genuchten



**Figure 26: Comparison of Leachate and water with Normalized  $\Theta_s$ ,  $\Theta_r$ , and dry densities (also showing non-normalized parameters)**

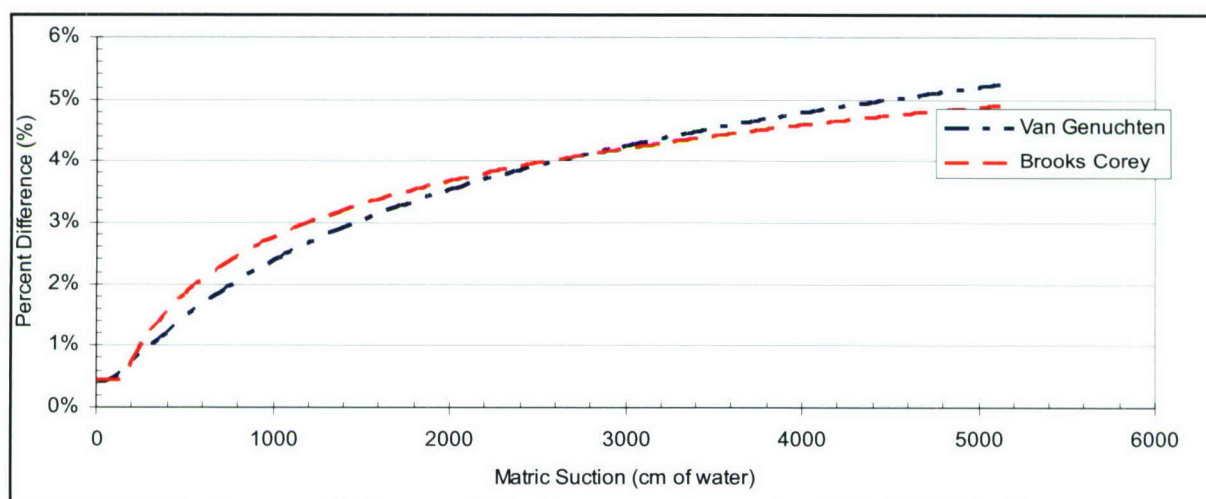


**Figure 27: Comparison of Leachate and water with Normalized  $\Theta_s$ ,  $\Theta_r$ , and dry densities (also showing non-normalized parameters)**

## 8 Discussions

### 8.1 Affects on the SWCC

Leachate has a discernable impact on the soil-water-characteristic curves. For both VG and BC models, the Leachate curve fell below the water curve. The percent difference between the leachate and water SWCCs is provided in Figure 28.



**Figure 28: Percent Difference between Leachate and Water Volumetric Water Contents**

Higher values of matric suction exhibit greater percent differences in computed or theoretical volumetric water contents. The range which is applicable for this research and for landfill liners in general lies below 2000cm of suction—this range corresponds to moisture contents at or above optimum which are appropriate for clay liners. For purposes relative to landfill liners, it appears leachate caused as much as a 25% difference in the SWCC.

The affects of leachate on individual characteristics such as air-entry value, steepness, inflection, and rotation are a bit more subtle. The most significant difference between models is how they handle the air-entry portion of the curve ( $\psi_a$ ), also called the bubbling pressure. Along a drying path, the SWCCs normally remain horizontal at saturated volumetric water content until  $\psi_a$  is reached at which point the curve begins to dip downward. While this point is not clearly defined on the Van Genuchten curve,  $\psi_a$  is distinctly represented on the BC SWCCs. The VG model depicts the curve as

immediately dropping away from saturation as matric suction increases. Although the VG model does not specifically define a variable for  $\psi_a$ , the parameter  $\alpha$  is commonly associated with the inverse of air-entry portion of the curve. Comparison of the results from the parameter optimization indicates that  $\alpha$  decreased by 4.0% when the soil was permeated with leachate, suggesting the air-entry increased. This is consistent to the BC model which indicates leachate increased the air-entry value from 149.1 to 158.0 cm, a 6.4 % increase. The affect on other SWCC parameters are provided in Table 18.

**Table 18: Leachate's Affect on BC and VG Parameters**

Model	Parameter	Parameter Optimization Results		% Difference
Van Genuchten	$\alpha$ (1/cm)	Water	0.003901	-4.0%
		Leachate	0.003747	
	n	Water	1.103105	2.4%
		Leachate	1.129887	
Brooks Corey	hb (cm)	Water	149.1409	6.4%
		Leachate	158.968	
	$\lambda$	Water	0.091387	18.2%
		Leachate	0.109697	

The other parameter in the Brooks Corey model is the pore-size distribution index,  $\lambda$ , which showed an 18.2 % decrease indicating a steeper slope about the pivot point. A decrease in the slope controlling parameter in the VG model,  $n$ , also indicates a steeper slope about the pivot point for the leachate SWCC.

## 8.2 Modeling Scenario #1

This scenario fixes the conductivity functions. The Brooks Corey and Van Genuchten Models both predict an initial increase in matric suction, hydraulic conductivity, and cumulative flux for the soil column permeated with leachate. This difference is more pronounced in the Van Genuchten models. The initial differences between water and leachate for the VG and BC SWCC models are summarized in the following tables.

**Table 19: Van Genuchten Difference in Initial Properties**

Property	Water	Leachate	Difference	% Difference
Initial Matric Suction (cm)	1086.99	718.252	368.738	-41%

	Property	Water	Leachate	Difference	% Difference
	Initial Hydraulic Conductivity (cm/hr)	9.73E-08	2.22E-07	-1.23E-07	78%

**Table 20: Brooks Corey Difference in Initial Properties**

	Property	Water	Leachate	Difference	% Difference
	Initial Matric Suction (cm)	914.378	690.385	-223.993	-28%
	Initial Hydraulic Conductivity (cm/hr)	1.38E-07	2.4E-07	-1.02E-07	54%

There is initially a relatively significant 78% and 54% difference in unsaturated conductivities at the start of the of the VG and BC analysis which can be contributed to an initial drop of 41% and 28%, respectively, in matric suction. Because the unsaturated conductivity function is dependent on matric potential (even though the conductivity functions themselves are fixed in this scenario), it would follow that a significant change in matric suction would cause a change in overall potential and thus a change in hydraulic conductivity. The difference in matric suction is derived from the difference in predicted potentials for the given initial volumetric water content; in this case 0.399 for both permeates. At this water content, the resulting soil water characteristic curves predict an initial matric potential of -1086.99 cm for water and -718.25 cm for leachate (1086.99 cm and 718.25 cm, respectively, when expressed as matric suction). Solving Richards' equation quantifies this matric potential difference into unsaturated hydraulic conductivity.

The most significant difference between the BC and VG models is the time lag between when permeates begin to saturate the soil at 61.0 cm. In both cases the leachate is the first to go toward saturation. The BC model predicts scarcely any time lag between saturation times for both permeates, while the VG model predicts almost a 200 hour lag between saturation times. For the VG SWCC model, saturation occurs between 5000 and 6000 hours after ponding and the two permeates do not become synchronized until well after saturation, while saturation in the BC model occurs between 6000 and 7000 hours and the permeates sync up relatively quickly. The VG model reaches saturation quicker than the BC model despite BC having a higher initial hydraulic conductivity for both permeates. This can be explained by how initially the BC SWCC curves lie below the VG curves—therefore predicting lower matric suction for the same water content and consequently a

higher hydraulic conductivity. However, as the water content increases the BC curves cross over the VG curves thus allowing the VG conductivity to increase faster relative to the BC conductivity. The two models handle the air suction portion of the SWCC very differently, and this is the reason for the curve cross over as water contents head toward saturation. Differences in the air-suction portion of the curve are explained in more detail in Section 8.1.

### 8.3 Modeling Scenario #2

This scenario allows for independent conductivity functions. The Brooks Corey and Van Genuchten Models both predict an initial increase in matric suction, hydraulic conductivity, and cumulative flux for the soil column permeated with leachate. The divergence of this scenario from the fixed conductivity scenario is some of the differences between the two permeates are more pronounced. It is expected that matric suction would not change with a change in permeability because it is a function of only the SWCC. The initial differences between water and leachate for the VG and BC SWCC models are summarized in the following tables.

**Table 21: Van Genuchten Difference in Initial Properties**

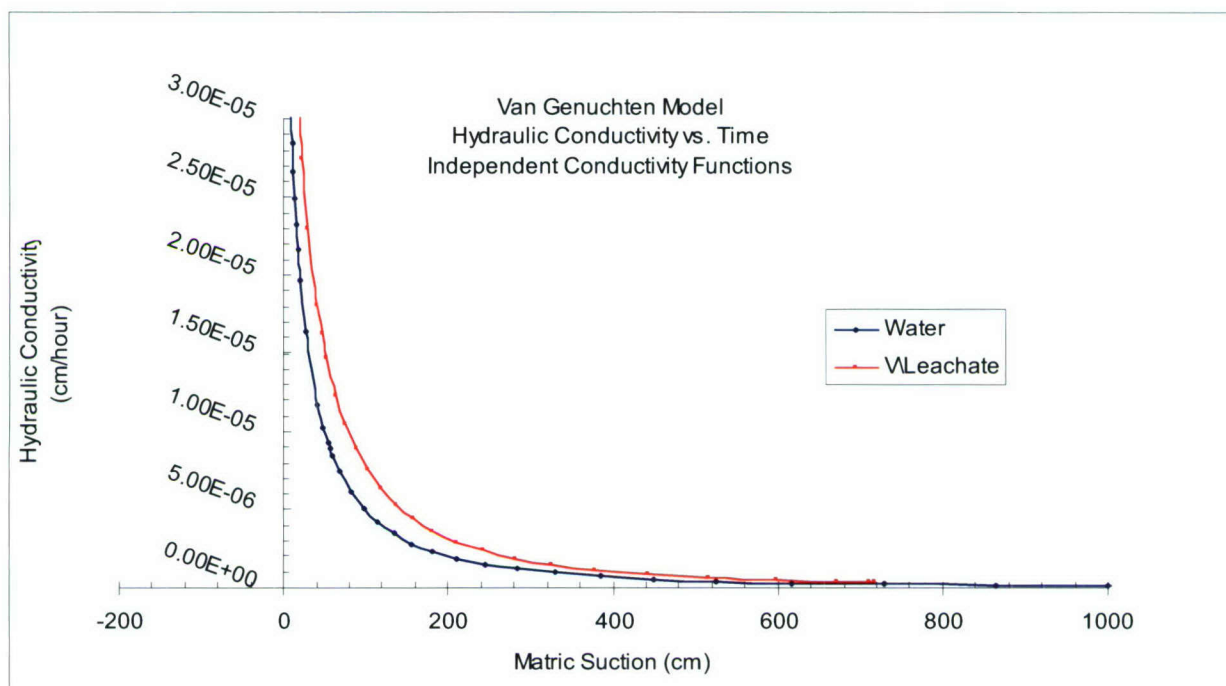
Property	Water	Leachate	Difference	% Difference
Initial Matric Suction (cm)	1086.99	718.252	368.738	-41%
Initial Hydraulic Conductivity (cm/hr)	9.73E-08	3.4E-07	2.43E-07	111%

**Table 22: Brooks Corey Difference in Initial Properties**

Property	Water	Leachate	Difference	% Difference
Initial Matric Suction (cm)	914.378	690.385	196.126	-28%
Initial Hydraulic Conductivity (cm/hr)	1.38E-07	3.67E-07	2.29E-07	91%

This increase in relative difference is expected because of the relationship between matric suction and permeability. A change in the SWCC will cause a proportional change in the conductivity equation. Whereas Scenario One showed only the difference which was the result of changing SWCCs, the differences in this scenario are compounded by the change in conductivity.

In terms of how hydraulic conductivity changes with respect to matric suction, it is of interest to look at how leachate is different from water. Figure 29 shows this relationship based on the Van Genuchten SWCC. It is important to note in this modeling scenario non-optimized conductivity parameters are used based exclusively upon the SWCC parameters. This is an accepted procedure for quantifying relative differences.



**Figure 29: Hydraulic Conductivity vs. Matric Suction**

#### 8.4 Modeling Scenario # 3

Part of the variation in the SWCC curves, and in turn hydraulic conductivity and cumulative flux, is due to experimental variation—in this case compaction conditions. The dissimilarity of the water and leachate soil water characteristic curves is the result of a difference in Van Genuchten parameters and saturated and residual water contents—these are the inputs to Richards' equation and also CHEMFLO. While the difference in VG Parameters are the result of varying equilibrium water contents as measured by the Pressure Plate Extractor, the saturated and residual volumetric water contents are derived from geotechnical phase relationships which are dependent upon average compacted dry density for the soil samples. Under ideal laboratory conditions, compacted dry density would be the same for samples permeated with leachate and water, but slight difference

in compaction conditions resulted in a different derived saturated and residual water contents. While the percent differences for average dry density vary by less than 0.4%, (0.00541 g/cm<sup>3</sup>) this variation caused a 5 to 10% increase in hydraulic conductivities and matric suction for both models.

Table 23 shows the difference between computed hydraulic conductivity and matric suction at  $t = 0$  and  $x = 61$  cm (the bottom of the soil column) for the two permeates with different dry densities and computed residual and saturated water contents. Table 24 shows the reduction in percent differences by fixing saturated and residual volumetric water contents – taken to be the average of the two.

**Table 23: Initial Hydraulic Conductivity and Matric Suction**

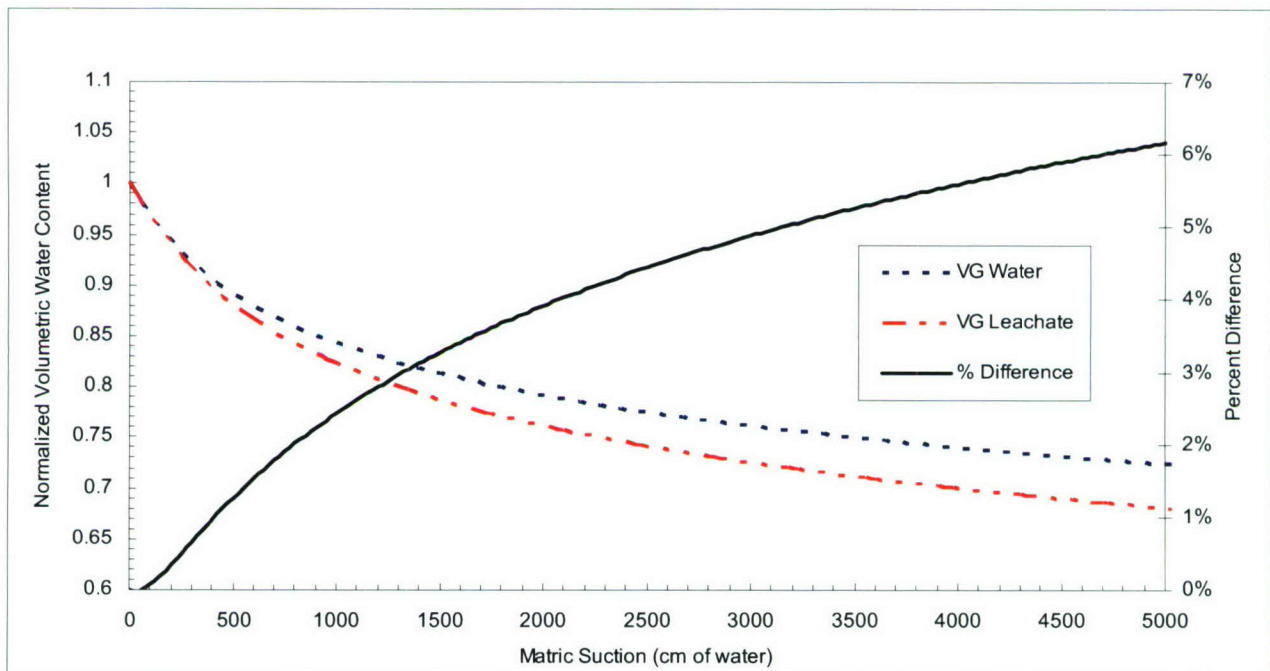
Property	Water	Leachate	Difference
Average Dry Density (g/cm <sup>3</sup> )	1.449324	1.454733	0.005409
Average Computed Volumetric Water Content (Saturated)	0.457556	0.455545	-0.002011
Average Computed Volumetric Water Content (Residual)	0.075229	0.075510	0.000281
BC Initial Matric Suction (cm)	-914.378	-690.385	223.993
BC Initial Hydraulic Conductivity (cm/hr)	1.38E-07	2.4E-07	1.20E-07
VG Initial Matric Suction (cm)	-1086.99	-718.252	368.738
VG Initial Hydraulic Conductivity (cm/hr)	9.73E-08	2.22E-07	1.247E-07

**Table 24: Normalized  $\Theta_s$ ,  $\Theta_r$ , and Dry Densities**

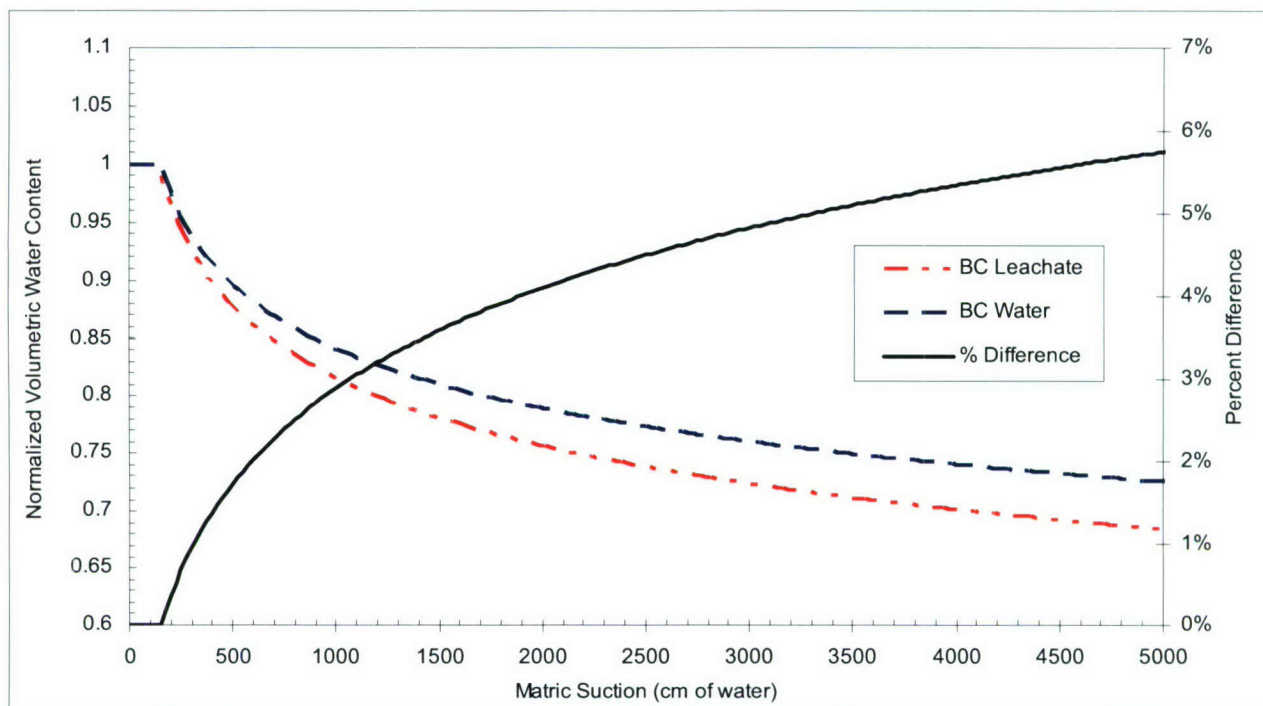
Property	Water	Leachate	Difference
Average Dry Density (g/cm <sup>3</sup> )	1.452029	1.452029	0
Average Computed Volumetric Water Content (Saturated)	0.456550	0.456550	0
Average Computed Volumetric Water Content (Residual)	0.075370	0.075370	0
BC Initial Matric Suction (cm)	-894.13	-706.787	187.343
BC Initial Hydraulic Conductivity (cm/hr)	1.44E-07	2.293E-07	8.53E-08
VG Initial Matric Suction (cm)	-1054.80	-737.36	317.44
VG Initial Hydraulic Conductivity (cm/hr)	1.04E-07	2.112E-07	1.072E-07

Studies conducted by Merayyan et. al. have shown compaction conditions play an important role in SWCC determination with varying compactions changing the curve-fit parameters by several orders of magnitude. It is therefore possible that even a small difference in compaction, in this case 0.4% could have an affect on the SWCC. In this study it is not possible to delineate the real effect of varying compaction conditions on the SWCC and thus hydraulic conductivities from the apparent affects caused by using theoretical relationships dependent on dry density to convert from gravimetric water contents to volumetric water contents. It is most likely a combination of these two phenomena that are responsible for an additional 5 to 10% difference between permeate hydraulic conductivities and matric suctions.

A common way to reduce the impacts of variations in  $\theta_s$ ,  $\theta_r$  on the SWCCs is to convert volumetric water content to normalized volumetric water content ( $\Theta$ ) with the relationship  $\Theta = (\theta_w - \theta_r) / (\theta_s - \theta_r)$ . This analysis is presented in Figure 30 and Figure 31. These percent differences are very close to those offered in Figure 28 for standard SWCC's.



**Figure 30: Van Genuchten Normalized Volumetric Water Content for Leachate and Water**



**Figure 31: Brooks Corey Normalized Volumetric Water Content for Leachate and Water**

## 9 Summary and Conclusions

Municipal landfill leachate and clay which was used in construction of a low permeability liner were used to assess the impacts which leachate could have on the soil-water-characteristic curve and subsequent flow characteristics. A pressure plate extractor was used in conjunction with the axis-translation technique to produce matric suction vs. water content data. The SWCC's were developed based on PPE data and curve fit to the Van Genuchten and Brooks Corey models using Excel solver routine for parameter optimization. A numerical analysis was conducted with CHEMFLO-2000 using a finite differences approach to solving Richards Equation for the change in matric suction, unsaturated conductivity, and cumulative flux with respect to time. Because the purpose of this research is to discuss the impacts leachate might have on the SWCC and consequent flow conditions through unsaturated clay as it pertains to landfill liners, every effort was made to model the soil in terms of realistic landfill conditions where feasible. Decisions about modeling inputs such as initial conditions, boundary conditions, and soil column properties are managed to match field conditions.

The research indicates leachate causes a decrease in matric suction for given moisture content and consequently an increase in hydraulic conductivity and cumulative flux

through a common landfill clay for both Van Genuchten and Brooks Corey models. The largest difference in predicted matric potentials for water contents between saturation and those appropriate for landfill liner design vary depending on the modeling scenario utilized. On average these differences ranged from zero to 209cm at saturation (20.5 kPa) for Brooks Corey and 368cm (36.1 kPa) at saturation for Van Genuchten. Hydraulic conductivities also increased when permeated with leachate. Increases in cumulative flux are mostly dependent on the time lag between when leachate and water begin saturation.

The Brooks Corey and Van Genuchten models handle the air-entry portion of the SWC quite differently and this had a substantial impact of the outcome of hydraulic conductivity vs matric suction relationships. Consequently any predictions of flow conditions which depended on this portion of the curve, below a matric potential of about 200 cm (19.6kPs) was heavily dependent upon the choice of model. Because the numerical modeling appropriate for landfills depended on this the air-entry portion of the curve, the change in hydraulic conductivity from water to leachate depended on the model. Even though there was a discernable difference in both models, the largest difference was not more than  $2.4 \text{ E-7 cm/hour}$ .

These findings are consistent with historical research detailing how pore fluid chemistry has a definite affect on flow characteristics. The most recent study, published by Ichola & Gaidi at the same time this research was being conducted, concluded leachate causes an increase in hydraulic conductivity when measured with methods similar to those in this study. Ichola & Gaidi found that even though permeability was influenced by leachate, the increase was not more than  $4.32 \text{ E-5 cm/hour}$ .

## **9.1 Engineering Implications**

Current regulations and common landfill liner design practices use a water-saturated permeability approach for specifying liner materials. While this approach utilizes a specification which can be easily tested in the laboratory, it does not represent true conditions in the field. This research has attempted to more closely model field conditions for the time between liner construction and saturation. A complete analysis of how leachate might flow through a landfill liner while also considering the chemical

affects by leachate on flow characteristics is a significant contribution of this research. Most previous analysis of chemical pore-fluid interactions and its affects on the soil water characteristic curve or flow characteristics have been in some part different from the unsteady-state approach used in this study (with the exception of the 2006 study by Ichola & Gaidi).

While it has been determined that leachate increases unsaturated permeability, although not more than  $1\text{E-}7$ , it cannot be concluded that current regulations for liner designs are deficient. From an environmental perspective, specifying intrinsic permeability may be a more conservative approach in specifying permeability characteristics; unsaturated permeability tends to be lower than intrinsic values for clay soils. Better understanding of desiccation cracking, wilting point, and field capacity of materials used for landfill liners and covers might stand to gain from understanding how design criteria might be impacted by changes in pore fluid chemistry. The bottom-line is pore fluid chemistry has a real affect on properties of clayey soils. Engineers should be aware of the limitations invoked by neglecting pore fluid chemistry.

## **9.2 Future Research**

While this research was conducted with the goal to develop a comprehensive analysis of affects of municipal leachate on the SWCC and fluid movement through unsaturated clay, considering every facet of such a complex process is difficult to accomplish in one study. Therefore additional research is recommended to develop more robust conclusions about these effects.

This study utilized a single, relatively low PI soil. Additional clays with varying degrees of plasticity indexes would broaden the applicability of results to include a wider range of clays used in landfill liner construction. Research by Merayyan et al. has shown the SWCC depends on compaction conditions. This research has utilized a standard compaction method to prepare samples. A different compaction method, like the modified proctor which better simulates how clay is compacted in the field by a sheepsfoot roller, would also improve applicability. Design of these liners and their ability to protect groundwater resources might be enhanced by assessing the affects of leachate on strength characteristics in clayey soils including shear strength and shrink-swell properties. The performance of other properties which are based on the SWCC like

desiccation cracking and wilting point could be improved by additional research which includes both sorption and desorption portions of the SWCC—hysteresis of the desaturation curve was neglected in this study.

Limitations of the numerical methods used in this analysis could be improved by research relating to multi-dimensional models, instead of the one-dimensional model used by CHEMFLO. However, numerical analysis is limited by the resolution of experimental data. The PPE was used to measure matric suctions ranging from  $\frac{1}{2}$  bars to 5 bars (510cm to 5100cm of water pressure), however, the range for which is applicable for landfill liners, from optimum moisture content to saturation, corresponded to less than 1.5 bars. Therefore more data points of matric suction and water content should be tested in this range to improve experimental resolution. The PPE is a robust and very accepted method for testing this relationship; however testing experimental data at tighter range might be difficult with the current methods. Mitchell and Madison concluded that permeability was highly dependent on the testing method; therefore future research should rely on a variety of methods which, if feasible, are capable of gathering data at higher resolutions.

## 10 References

- Babu, G.L. Sivakumar; Peter, J.; Mukesh, M.D.; Gartung, E. "Significance of Soil Suction and Soil Water Characteristic Curve Parameters" *Geotechnical Testing Journal* Volume 28 (2005) 102-107.
- Bagachi, A. *Design and Construction and Monitoring of Landfills* 2nd Edition. New York: 1994
- Bowders, John J.; Daniel, David E.; Broderick, Gregory P.; Lilj. "Methods for Testing the Compatibility of Clay Liners with Landfill Leachate." *ASTM Special Technical Publication* (1985) 233-250.
- Dunn, Jefferey and Singh, Udai. "Landfill Closures: Environmental Protection and Land Recovery." *Geotechnical Special Publication No. 53* (1995).
- Fang, H.Y. *Introduction to Environmental Geotechnology*. New York: 1997.
- Fredlund, D.G. and Rahardjo, H. *Soil Mechanics for Unsaturated Soils*. New York: 1993.
- Gannet Fleming Inc. *Composite Liner Construction Plant for the Chicago Grade Landfill: Modules 3,4,5. Engineers Report: August 2004.*
- Goldman et al. "Clay Liners for Waste Management Facilities." *Pollution Technology Review No. 178* (1990): 27-232.
- Hathhorn, Wade E. "Review of Fluid Movement in Compacted Clay: The Case of Macropore Flow." *Engineering Hydrology: Proceedings of the Symposium* (1993) 472-478.
- Hickman, H. Lanier Jr. and Eldredge, Richard W. "A Brief History of Solid Waste Management in the US During the Last 50 Years." *MSW Management*. 2000. [http://www.forester.net/msw\\_0001\\_history.html](http://www.forester.net/msw_0001_history.html).
- Ichola, Alimi and Gaidi, L. "Hydraulic Ocnductivity and Pollutant Dispersion Coefficient Assessment during Leachate Flow in Unsaturated Clay" *Geotechnical Special Publication No. 147 Vol 2* (2006) 1547-1558.

- Karvonen, T. Soil and Groundwater Hydrology. Helsinki University of Technology: 2001
- Khire, Milland V. et al. "Water Balance Modeling of Earthen Final Covers." Journal of Geotechnical and Geoenvironmental Engineering. (1997) 744-752.
- Koerner, Robert and Daniel, David. Final Covers for Solid Waste Landfills and Abandoned Dumps. Virginia: 1997.
- Leong, E.C. and Rahrdo, H. "Permeability Functions for Unsaturated Soils." Journal of Geotechnical and Geoenvironmental Engineering. (1997) 1118-1126.
- Lisboa, R and Azevedo, R, and Martins Reis, R "Residual Hydraulic Conductivity Determination using Field and Laboratory Methods Geotechnical Special Publication No. 147 Vol. 2 (2006) 1559- 1571.
- Miller, Carol J.; Merayyan, Saad; Yesiller, Nazli. "Impact of Soil Type and Compaction Conditions on the Soil-Water Characteristic Curve". Journal of Geotechnical and Geoenvironmental Engineering (2002): 733-742.
- Miller, Carol J.; Merayyan, Saad; Yesiller, Nazli. "Unsaturated performance comparison of compacted clay landfill liners". Geotechnical Special Publication No. 99 (2000): 555-568.
- Miller, Carol P.; Wright, Steven J. "Predicting Leakage through Clay Landfill Covers." National Conference on Environmental Engineering (1984) 708-710.
- Mitchell, James K.; Madsen, Fritz T. "Chemical Effects on Clay Hydraulic Conductivity." Geotechnical Special Publication No. 13 (1987) 87-116.
- Nofziger, D.L. and Wu, Jinquan "Interactive Software for Simulating Water and Chemical Movement in Unsaturated Soils." CHEMFLO-2000 User Reference (2003): 4-33.
- Oweis, Issas S. And Khera, Raj P. Geotechnology of Waste Management 2nd Edition. San Francisco: 1998.
- Pacific Waste Services, 2001 Report of Waste Discharge for Paso Robles City Landfill. Engineers Report to the Central Coast Regional Water Quality Control Board: March 2001.

- Pierce, J. Jeffrey; Salfors, Goran; Peel, Thomas A.; Witter, Kelly A." Effects of Organic Leachate on Select Clays." Journal of Geotechnical Engineering Vol 113 (1987) 915-919.
- Quigley, Robert M et al. "Hydraulic Conductivity and Diffusion Monitoring of the Keele Valley Landfill." Canadian Geotechnical Journal Vol 30 (1993): 124-134.
- Quigley, Robert M et al. "Hydraulic Conductivity of Contaminated Natural Clay Directly Below a Domestic Landfill." Canadian Geotechnical Journal Vol 24 (1987): 377-383.
- Quigley, Robert M.; Rowe, R. Kerry. "Leachate Migration through Clay Below a Domestic Waste Landfill." ASTM Special Publication (1986): 93-103.
- Simpson, Brian E. "Evaluation of Leachate Compatibility to Clay Soil for Three Geosynthetic Clay Liner Products." Geotechnical Special Publication No. 103 (2000): 117-128.

## **Evaluating the effects of aging on liquefiable soil deposits**

Project Investigators:

Robb Moss and Gregg Fiegel  
Civil and Environmental Engineering  
California Polytechnic State University  
San Luis Obispo, California

## ***Summary***

This is the final report on the C3RP Project “Evaluating the Effects of Aging on Liquefiable Soil Deposits.” This was a pilot project to demonstrate the viability of performing laboratory testing on sands in order to determine the effects of geologic aging on liquefaction resistance. This pilot project succeeded in; a) the site selection and planning of the field component of the research, b) staffing of the project with undergraduate researchers, c) installation, training, and testing with the new lab equipment, and d) producing initial results for subsequent analysis. This project has presented many interesting but surmountable obstacles, and future multi-year external funding for this project looks promising.

## ***Introduction***

Liquefaction, an earthquake phenomenon where saturated sandy soil loses significant strength due to ground shaking, caused significant earthquake related damage observed during the Loma Prieta and Northridge earthquakes. The potential for soil liquefaction is commonly assessed using empirical methods that are based on post-earthquake field case histories of liquefaction. Case histories are generally preferred to laboratory tests because acquiring “undisturbed” samples of liquefiable material for lab testing can be difficult.

The existing database of liquefaction case histories is currently limited to soil deposits of late-Holocene age (within the last few thousand years). In general, soils older than late-Holocene have shown higher resistance to liquefaction due to aging effects (e.g. “cementation”). To date, some qualitative increases in liquefaction resistance due to aging have been documented, but there are no analysis methods currently available that quantitatively account for aging effects. As a result, it is difficult to accurately determine the liquefaction potential of soils older than late-Holocene in age. This poses a significant hazard for the metropolitan areas of San Francisco and Los Angeles where portions of the built environment are located on various soil deposits older than late-Holocene.

We propose to evaluate the effects of aging on soil liquefaction resistance by testing soil deposits older than late-Holocene. Aged deposits of liquefiable material generally exhibit some “cementation” thereby making these deposits better candidates for “undisturbed” block sampling. For this project we retrieve and test block soil samples which permit the use of both field and laboratory test results during our liquefaction evaluations.

## ***Background***

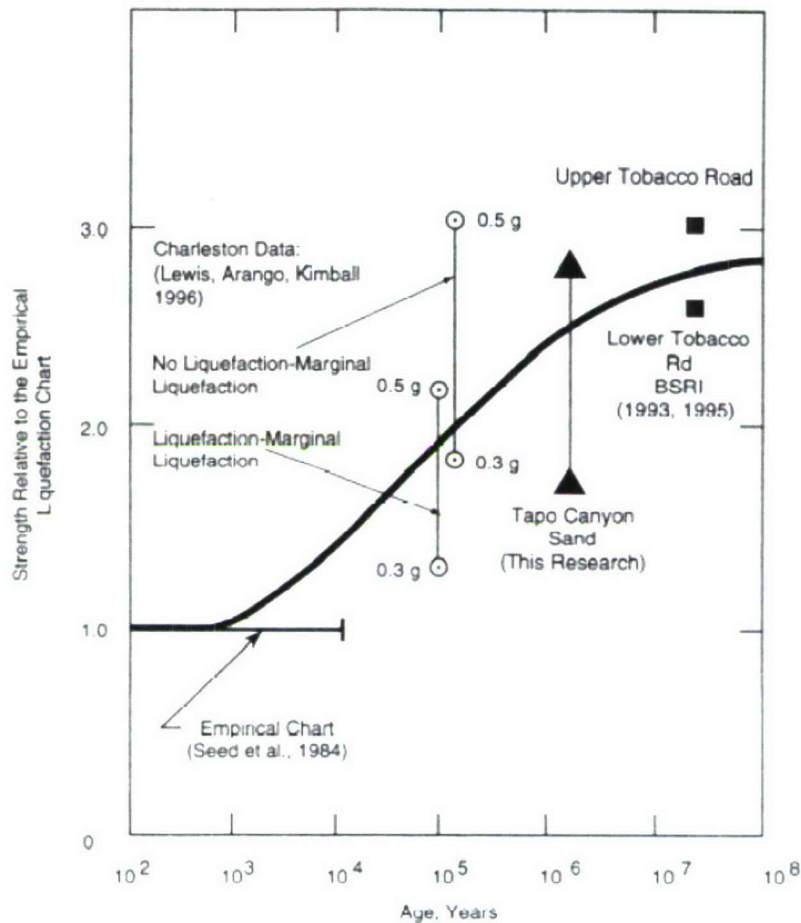
Current methods for estimating liquefaction potential are limited to soil deposited in the last few thousand years (Moss et al. 2006). Previous researchers have noted that time dependent changes in soil properties, or aging effects, can have a significant influence on liquefaction resistance. Aging effects have been observed in both short-term ( $10^{-2}$  to  $10^1$  years) and long-term ( $10^1$  to  $10^6$  years) time scales. Yet, aging effects are neglected in the current state of practice of liquefaction evaluation. As stated in the summary report of the recent NCEER workshop on liquefaction (Youd et al. 2001), “Although qualitative

increases in liquefaction resistance have been well documented, insufficient quantitative data have been assembled from which correction factors for age can be defined.”

Baxter and Mitchell (2004) have shown that all manner of tests to mimic the aging process in a laboratory environment fail to systematically capture the phenomena that cause aging effects (e.g. cementation, soil fabric, influence of pore fluid chemistry, particle bonding, etc.). On the other hand, in situ field measurements made using the cone penetration test (CPT) have been shown to detect the effects of aging on penetration resistance (e.g., Mitchell and Solymar, 1984; Hryciw, 1986; Schmertman et al., 1986; Charlie et al., 1992). These results, however, have yet to be quantified in terms of liquefaction resistance.

Kulhawy and Mayne (1990) accumulated data from the literature showing the influence of aging on soil resistance. The results were presented as a statistical regression of the mean trend. This study is only a rough guide to the effects of aging because of the large data scatter from many different and not necessarily compatible tests.

Arrango and Miguez (1996) studied the effects of aging by testing Pleistocene-age, lightly cemented liquefiable material that was obtained from Tapo Canyon, CA. This site was selected because it was adjacent to a site underlain by Holocene-age soils that had liquefied during the 1994 Northridge earthquake. Arrango and Miguez tested soil samples that were carefully hand-carved from soil blocks in order to preserve soil fabric and cementation. Cyclic triaxial (CTX) laboratory tests were performed on hand-carved (i.e., “undisturbed”) and reconstituted soil samples to evaluate liquefaction resistance. Figure 1 shows the effects of aging on the Tapo Canyon sand with respect to other similar research. This study provided useful insights on aging, but focused on only one soil deposit.



**Figure 1. From Arrango and Miguez (1996), showing the interpreted affects of aging on soil liquefaction. The y-axis is the increase in cyclic strength due to aging, reported as a factor increase over the cyclic strength of Holocene-age soils.**

An objective of our research will be to expand upon the work of Arrango and Miguez (1996) by:

- Increasing the number of aged soil deposits evaluated;
- Using a laboratory test more appropriate than the cyclic triaxial test to evaluate liquefaction resistance (i.e. the Cyclic Simple Shear (CSS) test); and
- Estimating the reconstituted CPT resistance using a theoretical (i.e. cavity expansion) model.

### ***Research Methodology***

The primary objective of this research will be to systematically evaluate how aging can influence the liquefiability of a soil deposit with respect to the commonly used empirical liquefaction methods. This research will provide an age-based correction factor that can be applied during liquefaction resistance evaluations thereby eliminating guesswork as to how aging has impacted the liquefiability of an older deposit.

The following is a general list of steps needed to perform this research;

- At each site a block sample of the aged soil is acquired for laboratory testing.
- Field measurements of the in situ soil resistance is made using the Cone Penetration Test.
- In the lab Cyclic Simple Shear (CSS) tests are performed on hand carved samples to estimate the liquefaction resistance of the aged soil deposit.
- CSS tests are also performed on reconstituted samples to evaluate the difference in liquefaction resistance between the aged soil and a freshly deposited soil of the same source material.
- Estimates of CPT resistance for the reconstituted soil will be made using a theoretical model based on cavity expansion theory. The model will be calibrated using soil properties measured in the lab.
- The field measurements of the CPT resistance of the aged soil will then be compared to the estimated CPT resistance of the reconstituted soil.
- By matching the CPT resistance measurements to the CSS results we will arrive at an indexed measure of the change in liquefiability with age.

Site selection is based on the following criteria; the site is located in areas with high hazard potential (i.e. where critical facilities, dense population centers, and/or transportation corridors exist), is underlain by potentially liquefiable soils exhibiting a wide age range, generally from late-Holocene to early-Pleistocene ( $10^1$  to  $10^6$  years), and is accessible and amenable to the proposed field testing and block sampling.

Cone penetration tests (CPTs) will be performed at the selected sites to measure the in situ index properties of the deposit. The collection of field test data will allow us to correlate our results with the existing empirical liquefaction methods. Gregg Drilling/Gregg In Situ is contributing approximately \$7000 of in-kind field investigation work to the project. This contribution includes \$500 for mobilization of CPT equipment to each site, profession operation of the testing equipment, collection of CPT index measurements per ASTM standards at roughly \$7-10 per foot of exploration, and subsurface data in digital and print format for post-processing

This research specifically addresses three of the research focus areas described in the C3RP technology park initiative: seismic research, disaster management and mitigation, and GIS applications in engineering.

### ***Site Selection***

A site for this project must have potentially liquefiable soils that are older than recent-Holocene and are sufficiently “cemented” that a block sample can be acquired.

Potentially liquefiable means that a soil is primarily sand with a relatively low percentage of fines (soils smaller than sand). Optimally the soil would be composed of sand with less than 5% to 10% fines by mass, thereby minimizing the effects of fines content on the liquefaction resistance of the soil.

Age dating of soils can be a rather difficult and usually requires several techniques applied by a specialist in soil chronology and dating methods. For this project we pursued sites where the soils have been previously dated to avoid tackling the age dating ourselves.

Having a “cemented” sample is a must for this research. It is necessary in getting a sample from the field and into the testing device in a relatively undisturbed state, thereby mimicking the *in situ* resistance to cyclic loading and pore pressure generation. The term “cemented” is in quotes because there exist a number of reasons for sand to exhibit grain to grain bonding, but in this case we are referring to this bonding generically as a function of the aging process without delving into the bonding mechanism in detail.

Given the above criteria we proceeded to review potential sites and focus on the most likely candidates. The proposed research was discussed with various geo-professionals, and sites were recommended by Kieth Knudsen (CGS), Tom Blake (Fugro West), Ron Richman (CalTrans), Ignacio Arrango (URS), and Jon Blanchard (Fugro West). Below is a list of potential sites that were considered for this pilot project;

- 1) Antioch/Oakley dunes (Northern California)
- 2) San Benito River (Northern California)
- 3) Sunol quarry (Northern California)
- 4) Aromas formation N. Monterey county (Central California)
- 5) Old Santa Cruz quarry (Central California)
- 6) Santa Cruz beach terraces (Central California)
- 7) Los Osos dunes (Central California)
- 8) Nipomo Mesa (Central California)
- 9) Orcutt formation in Santa Barbara County (Central California)
- 10) Terrace deposits in Santa Barbara (Central California)
- 11) Manning Pit (Southern California)
- 12) Tapo Canyon (Southern California)
- 13) Grimes and Gillibrand Canyon (Southern California)
- 14) Sherman Oaks (Southern California)
- 15) Culver City-Ladero Heights (Southern California)
- 16) LA River near CalState Northridge (Southern California)

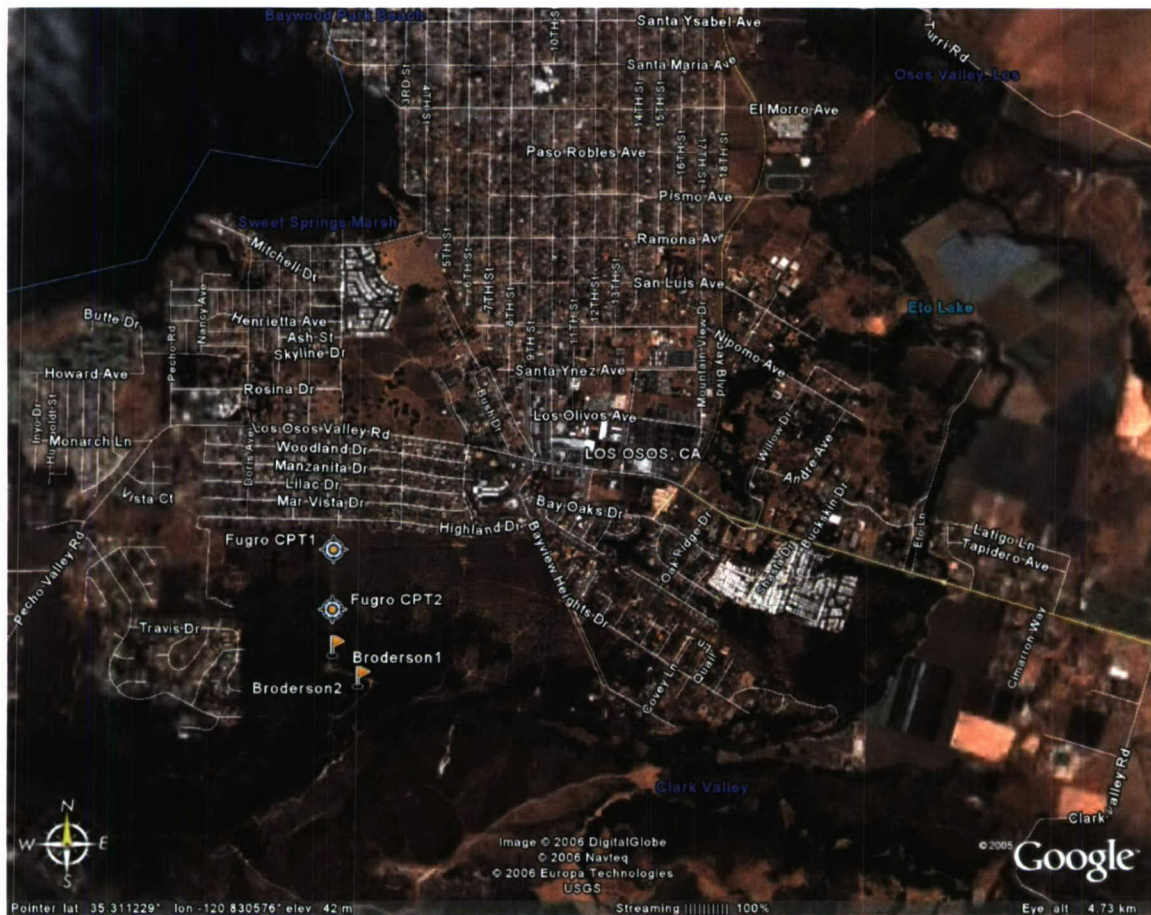
Of these, the sites that are currently known to contain “cemented” sands that have also been age-dated and are readily accessible for sampling are; Los Osos dunes, Tapo Canyon, and Manning Pit.

For the Los Osos site, the subsurface conditions and the age of the deposits have been reported by Jon Blanchard of Fugro West in a geotechnical report for the site. The proximity of the site to CalPoly and the good subsurface documentation made this site an excellent choice for this pilot project. Several block samples were acquired and have been subsequently tested. Below is a map (Figure 2) showing the location of where the block samples were acquired.

Tapo Canyon was originally tested by (Arrango and Miguez 1996) for the effects of soil aging on liquefaction potential under a grant from the NSF. We contacted the author of the lab testing portion of the Arrango and Miguez study, Mike Riemer (UC Berkeley), and found that two block samples still existed in storage. Dr. Riemer graciously provided these block samples for our current testing. The landowner at the quarry where the block samples were acquired was contacted about our current project. To date the owners have been unreceptive to allowing us access to the site for subsequent field testing.

Manning Pit is a quarry in Irwindale, CA, that contains “cemented” sands. A report on testing of these “cemented” sands was provided by Tom Blake (Fugro West, personal communication 2006). The city of Irwindale has been contacted and negotiations for site access are underway.

The three sites (Los Osos, Tapo Canyon, and Manning Pit) will be explored further and block samples and subsurface investigations will be performed as logistics permit. These sites will also be documented in a GIS database that contains locations of “cemented” deposits, locations where samples were acquired, and locations of nearby CPT soundings.



**Figure 2. Location of Los Osos site where "cemented" samples were acquired. Block samples of two different units were acquired at Broderson1 and Broderson2. Also shown are nearby CPT sounding locations CPT1 and CPT2 from Fugro.**

### ***Student Researchers***

The C3RP proposal originally budgeted for one graduate student and two undergraduate researchers throughout the summer. There were not any graduate students readily available for this project so I targeted this project at undergraduate researchers.

The goal of lab testing using undergraduate researchers is to have students trained on the testing equipment so that they can; a) learn about soil mechanics using state-of-the-art geotechnical testing equipment, b) aid in the progress of existing research projects that utilize the testing equipment, c) assist in training new undergraduate researchers as they join the group, and d) potentially continue here or at other institutions in pursuit of an advanced degree.

In Winter quarter 2006, I made a presentation at several student group meetings to recruit undergraduate researchers for this, and other research projects I currently have underway. For this project I hired two highly motivated undergraduate researchers and we began working in the lab at the start of Spring quarter. They worked on the liquefaction project part time (5 to 8 hrs a week) during the quarter and near full time in the summer.

The lack of a graduate students had the advantage that it allowed me to hire more undergraduate researchers to work on this and other projects. With the graduate summer salary for this project and funding from other sources I have hired three other undergraduate researchers to work on a burgeoning project dealing with seismic levee stability in the California Bay Delta. This new project involves similar lab testing using the equipment in the advanced soils lab. The new hires are currently training with and assisting the existing researchers thereby accelerating the pace of both projects.

We are striving for a continuous flow of undergraduate researchers that are trained on the testing equipment and then transfer that knowledge to new researchers. We hope to find a select few who will continue on as graduate researchers here and then can assist in managing the undergraduate researchers. The C3RP funding has enabled us to start this flow of researchers and has allowed us to establish a lab system that we can rely on as we chase after additional funding for this and other lab based research projects.

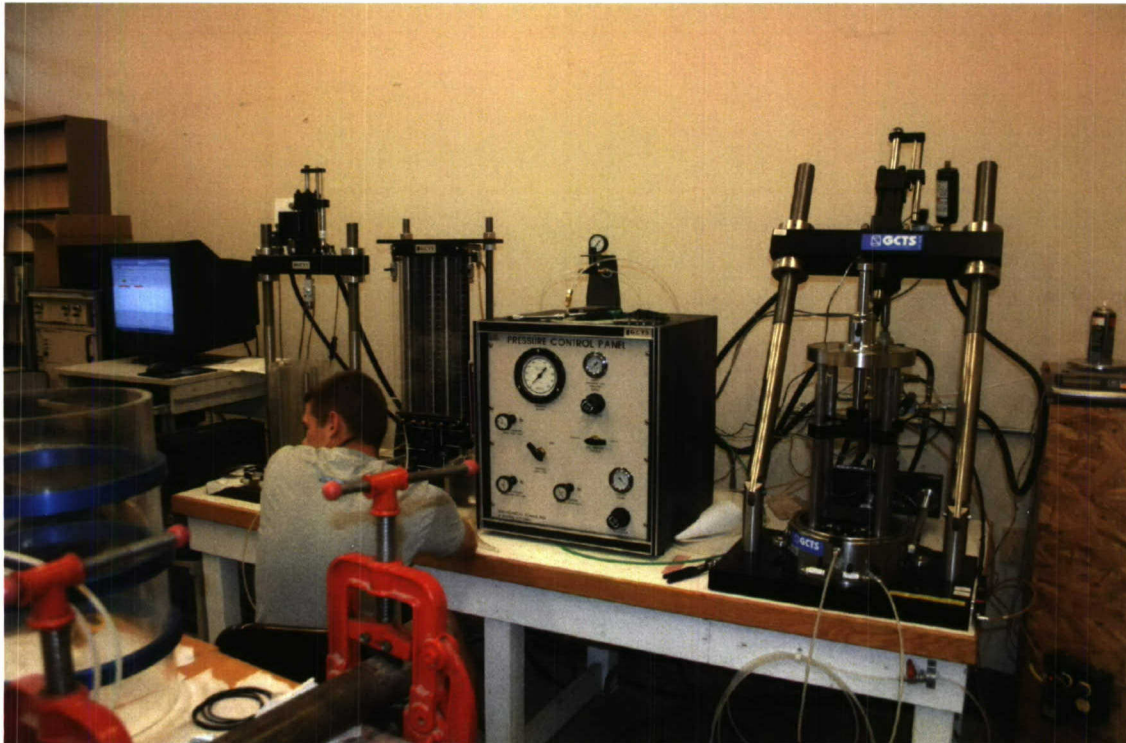
### ***Lab Testing Equipment, Training, and Testing***

In order to perform liquefaction testing on the cemented and reconstituted sand samples we had to set up the appropriate testing equipment in the advanced soils lab. A Cyclic Simple Shear (CSS) device (Figure 3) was purchased in the spring of 2005 (using a grant from the Student Fee Initiative) with the intent of performing testing as outlined in the C3RP proposal. This device was built by the testing equipment manufacturer GCTS, and delivered to CalPoly at the soonest possible date which was around the end of January 2006. Several days were dedicated to the installation and calibration of the new CSS device under the supervision of a GCTS technician.

Laboratory testing in general and liquefaction testing in particular is a non-trivial process. On top of that, we had the task of learning how to operate and optimize the new lab testing equipment. Our goal for Spring quarter was to; 1) familiarize ourselves with the

testing device, hydraulic system, control system, and operating software, 2) work out any bugs or glitches in the new system, and 3) become familiar with the many steps involved in running a standard liquefaction test using the CSS device.

There were several hardware and software bugs that took essentially half of the quarter to resolve. We are currently using one hydraulic and control system to run both the CSS device and another existing device (Triax). This was the most efficient and cost-effective means of equipping the lab for several types of soil tests. In order to allow for easy switching between the Triax and the CSS we had to install and calibrate two new A to D (Analog to Digital) cards. This process took several weeks and is an example of the type of equipment setup hurdles that we had to overcome in the spring.



**Figure 3. CSS device (on the far right) with the cell wall removed awaiting a new sample for testing.**

The two researchers who worked on the equipment setup and establishing testing procedures are now training three new researchers who will be using the same system on another project. This transfer of knowledge has proved to be a very effective means of educating researchers so far and we hope to continue this indefinitely.

Once we felt comfortable with the equipment and our skills at running a standard liquefaction test, we brought in an expert to provide us with some of the finer details on liquefaction testing. Dr. Annie Kammerer paid us a two day visit at the beginning of June to look at our equipment, observe our testing procedures, and provide some insight into the nuances of liquefaction testing. Dr. Kammerer earned a Ph.D. doing sophisticated liquefaction testing and has years of experience performing similar tests.

This visit was invaluable in the amount of information and insight she could offer us on aspects of sample preparation, equipment operation, and testing procedures. In attendance were myself and the co-PI as well as five undergraduate researchers at an all day lab meeting where we ran tests with Dr. Kammerer looking on and providing guidance.

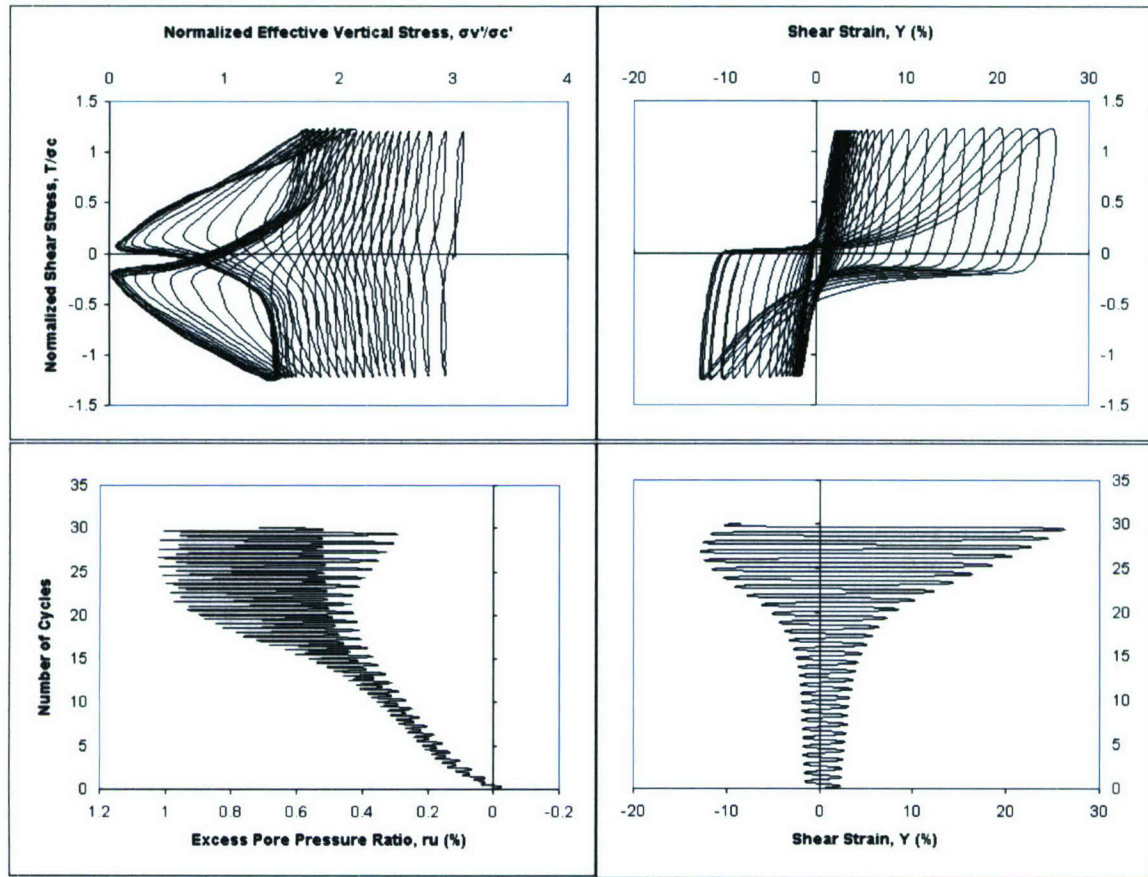
Our cyclic simple shear testing procedures and protocols closely follow the work by Kammerer et al. (2001) and Wu et al. (2003). Where we have diverged from the previous research is when addressing specifics that pertain to “cemented” sands. Additional guidance is provided by Dr. Mike Riemer at U.C. Berkeley who has previous experience testing “cemented” sands.

Upon working out the testing details and after successfully completing a number of tests, we began production mode. To document all the knowledge gained through the learning phase we recorded each procedure on a DVD, which will serve for future reference as well as for training new researchers. This DVD is currently 30 min long with a number of chapters that cover every phase of sample preparation, equipment preparation, and subsequent testing.

### ***Initial Test Results***

In production mode we are evaluating the liquefaction resistance of a “cemented” sample, and then evaluating the liquefaction resistance of a complimentary reconstituted sample. Because of the natural variability of soil deposits we must run multiple tests to determine a median liquefaction resistance for both the “cemented” and reconstituted samples of any one material. Show below in Figure 4 are typical results for one liquefaction test.

The preparation of a sample for testing can take anywhere from 4 to 10 hrs depending on the rate at which back saturation achieves a B-value of 0.96. The test itself generally lasts approximately 3 min. Figure 4 shows the complete test results. This plot will provide one data point based on the relative age of the material and the number of cycles to achieve liquefaction. Liquefaction here is defined as 5% single amplitude shear strain, or equivalently 2.5% double amplitude shear strain. A number of tests will be run on the “cemented” and reconstituted specimens to generate a pair of CSR versus number of cycles to liquefaction curves for this particular material. The resulting difference between these two curves will provide an equivalent magnitude specific measurement of the increase of liquefaction resistance as a function of the geologic age of the material. The goal is to test as many materials of different age that we can acquire to document the systematic change of liquefiability.



**Figure 4.** CSS test results for a “cemented” sand deposit. These plots from upper left in a counter-clockwise direction; effective vertical stress versus normalized shear stress, shear strain versus normalized shear stress, shear strain versus number of cycles, and excess pore pressure ratio versus number of cycles. This test was on a Los Osos tan specimen, CSR 0.6,  $\sigma_{v0}'=150$  kPa.

### *Future Funding Opportunities*

There is currently one pending federal proposal which would fund the continuation of this project. This proposal is a NSF CAREER proposal in which I used this research as one of four separate career goal projects that I hope to acquire funding for and perform research on. If this pending proposal is not funded a second more specific proposal will be submitted that targets this project in particular. The funding avenue that is most viable for this will be the RUI (Research at a primarily Undergraduate Institution) program within NSF. The fact that this project is currently producing successful results with undergraduate researchers should demonstrate the viability of this work and lend weight to the proposal.

### *Conclusion*

The liquefaction potential of aged deposits is a problem that currently has only been qualitatively addressed by previous researchers. Geotechnical consultants deal with the problem of aged deposits often enough that this issue is frequently discussed at workshops and conferences. Quantifying the effects of aging will result in better seismic

design and decreased seismic hazard for sites located on aged deposits. A well-researched correction factor for soil aging will readily be adopted by the consulting industry as well as state and federal agencies that must assess liquefaction potential.

This project will provide the field and laboratory results necessary for at least one journal publication as the research progresses and we gather more data. The testing methods employed in this study will be described in a separate journal publication. The results and testing methods would also be topical material for presenting at conferences and workshops.

### **References**

- Arrango, I., and Miguez, R. E. (1996). "Investigation on the Seismic Liquefaction of Old Sand Deposits."
- Baxter, C. D. P., and Mitchell, J. K. (2004). "Experimental Study on the Aging of Sands." *Journal of Geotechnical and Geoenvironmental Engineering*, 130(10).
- Charlie, W. A., Rwebyogo, M. F. J., and Doehring, D. O. (1992). "Time-Dependent Cone Penetration Resistance Due to Blasting." *Journal of Geotechnical Engineering*, 118(8), 1200-1215.
- Hryciw, R. D. (1986). "A study of the Physical and Chemical Aspects of Blast Densification of Sand," Ph.D. Dissertation, Northwestern University.
- Kammerer, A. M., Wu, J., Riemer, M. F., Pestana, J. M., and Seed, R. B. "Use of Cyclic Simple Shear Testing in Evaluation of the Deformation Potential of Liquefiable Soils." *4th International Conference on Recent Advances in Geotechnical Earthquake Engineering*, San Diego.
- Kulhawy, F. H., and Mayne, P. W. (1990). "Manual on Estimating Soil Properties for Foundation Design." *EPRI Report EL-6800*.
- Mitchell, J. K., and Solymar, Z. V. (1984). "Time-Dependent Strength Gain in Freshly Deposited or Densified Sand." *Journal of Geotechnical Engineering*, 110(11), 1559-1576.
- Moss, R. E. S., Seed, R. B., Kayen, R. E., Stewart, J. P., Der Kiureghian, A., and Cetin, K. O. (2006). "Probabilistic Seismic Soil Liquefaction Triggering Using the CPT." *Journal of Geotechnical and Geoenvironmental Engineering*, 132(8).
- Schmertman, J. H., Baker, W., Gupta, R., and Kessler, K. (1986). "CPT/DMT QC of Ground Modification at a Power Plant." *Use of In Situ Tests in Geotechnical Engineering*, ASCE STP No. 6.
- Wu, J., Seed, R. B., and Pestana, J. M. (2003). "Liquefaction Triggering and Post Liquefaction Deformation of Monterey 0/30 Sand Under Uni-Directional Cyclic Simple Shear Loading." *Geotechnical Engineering Report No. UCB/GE-2003/01*.
- Youd, T. L., Idriss, I. M., Andrus, R. D., Arango, I., Castro, G., Christian, J. T., Dobry, R., Finn, W. D. L., Harder, L. F., Hynes, M. E., Ishihara, K., Koester, J. P., Liao, S. S. C., Marcuson, W. F., III, Martin, G. R., Mitchell, J. K., Moriwaki, Y., Power, M. S., Robertson, P. K., Seed, R. B., and Stokoe, K. H., II. (2001). "Liquefaction Resistance of Soils: Summary Report from the 1996 NCEER and 1998 NCEER/NSF Workshops on Evaluation of Liquefaction Resistance of Soils." *Journal of Geotechnical and Geoenvironmental Engineering*, 127(10).

## Publications and Presentations Related to Project

### *Abstracts*

*Moss; Robb Eric S.; 'Preliminary Investigations into the Influence of Geologic Aging on Liquefaction Potential'; Presentation for Geotechnical Earthquake Engineering and Soil Dynamics Conference; May, 2008.*

This paper describes laboratory-based liquefaction testing on aged soil deposits to address the need for quantifying the influence of geologic aging on liquefaction potential. The in situ aging of liquefiable soil deposits has been qualitatively shown to result in an increase in liquefaction resistance. For example, Pleistocene age sandy soils will generally have a higher resistance to seismic induced liquefaction than similar sandy soils of Holocene age. However, a quantitative measure of this liquefaction resistance increase for use in liquefaction analysis is still unrealized. Cyclic simple shear testing was used to systematically quantify the influence of geologic aging on liquefaction potential for a number of aged soil deposits. The influence of geologic aging is then measured as the difference between liquefaction triggering of an aged sample versus a similar reconstituted sample. In situ resistance measurements of aged deposit were made using the Cone Penetration Test. Similar estimates of cone tip resistance of reconstituted samples were made using a cavity expansion model. These resistance values are used to tie the laboratory results to existing CPT-based liquefaction triggering correlations. Uncertainty in the field and lab testing results is explicitly defined in a probabilistic manner so that the age-based correction factor can be used in performance-based analysis. Preliminary results from this ongoing research project are presented.

## **Development of a hybrid rocket motor facility for advanced nozzle research**

Project Investigators:

William R. Murray and Thomas W. Carpenter  
Mechanical Engineering  
California Polytechnic State University  
San Luis Obispo, California

## Summary

Significant externally funded research into thrust vectoring aerospike rocket nozzles is in progress at Cal Poly. In this research program, aerospike nozzles are designed and built, and subsequently tested in the large cold-flow, blow-down facility in the ME Department. The aerospike nozzle offers the significant advantage of altitude compensation over conventional bell-shaped nozzles. With a movable spike, radial, longitudinal, and angular spike displacements from the nominal spike position induce changes in nozzle flow, which may provide efficient thrust vectoring and throttling for the annular aerospike nozzle. Results to date are quite promising for nozzle supply pressures up to the 120 psia level that is achievable with the current blow-down facility, but also show the need for testing nozzles at pressures around 300 psia. Under NASA funding, our cold-flow facility is being redesigned and upgraded to provide nozzle supply pressures of 300 psia.

An alternative way to test aerospike nozzles is to use the hot flow of high pressure gas generated by a hybrid rocket motor. Hybrid rocket motors are inherently safe because both the fuel and oxidizer can be stored without special precautions, and they do not burn when brought together without a significant source of heat for ignition. The development of a hybrid rocket motor facility for the testing of advanced rocket nozzles is the objective of this C<sup>3</sup>RP project. In this project, the existing Cal Poly hybrid rocket motor facility has been redesigned and upgraded to state-of-the-art research standards. The ultimate propellant combination that has been chosen for this hybrid is liquid nitrous oxide (N<sub>2</sub>O) as the oxidizer and hydroxy-terminated polybutadiene (HTPB) as the fuel, however both HTPB and Plexiglas have been used as the fuel in tests to date. For this hot flow environment, conical rocket nozzles have been fabricated from high-density graphite, however, ultimately aerospike nozzles will be fabricated from high-density graphite, carbon matrix composite materials, superalloys, or combinations of these materials. At the end of the first year of this project, twenty-two successful hot-fire tests have been conducted, including tests in which peak chamber pressures of 450 psia and average chamber pressures of 325 psia have been recorded. However, to prolong nozzle life, more work is needed in designing the geometry of the fuel grain such that a sustained chamber pressure of 300 psia can be achieved with an oxidizer/fuel mixture ratio near stoichiometric.

Not only will this facility be beneficial for the testing of rocket nozzles, but the marriage of hybrid rocket motors and thrust vectored aerospike nozzles is interesting in its own right. The potential for hybrid rocket motors was clearly demonstrated in 2004, when SpaceShipOne won the Ansari X-Prize using an HTPB-N<sub>2</sub>O hybrid rocket. In the future, hybrid rocket motors equipped with aerospike nozzles with thrust vectoring and throttling capabilities could play a very important role in programs trying to achieve inexpensive access to low earth orbit for NASA and the commercial sector, as well as becoming an inexpensive alternative for tactical missiles having throttling capabilities.

## Project Significance

This C<sup>3</sup>RP project is an important part of a growing research program in advanced rocket nozzle research being carried out in the Mechanical Engineering Department at Cal Poly. Accordingly, the broader research program and its significance will be described first, and then the significance of this C<sup>3</sup>RP project will be explained in that larger context.

### *Significance of the Advanced Rocket Nozzle Research Program*

Profs. Tom Carpenter and Bill Murray, along with Mr. Jim Gerhardt of the ME Department Technical Staff, who between them have decades of thrust vectoring experience, have recently been applying this expertise to develop advanced rocket nozzles with thrust vectoring capabilities. Attention has centered on the annular aerospike nozzle. This nozzle offers a significant advantage over conventional bell-shaped nozzles: altitude compensation, that is, the aerospike nozzle yields near optimal thrust at all altitudes, whereas a bell-shaped nozzle provides optimal thrust only at its design altitude. With a movable spike, radial, longitudinal, and angular spike displacements from the nominal spike position induce changes in nozzle flow, which may provide efficient thrust vectoring and throttling for the annular aerospike nozzle.

The optimal thrust vectoring and throttling of an annular aerospike nozzle are being investigated for solid rocket motors in a current NASA STTR<sup>1</sup>. In this NASA STTR project, scale models of nozzles are designed, built, mounted to a six-component thrust stand, and then subjected to cold air, blow-down tests. During the blow down, force data and color Schlieren videos are recorded. The force data are processed to provide the performance of the model nozzles in terms of thrust, pitch and yaw, and the color Schlieren images provide an in-depth insight into the resultant nozzle flow phenomena.

For this NASA STTR project, various nozzle configurations are being fabricated and tested to determine the optimum configuration for thrust vectoring. The first six aerospike nozzle prototypes, which were fabricated on a rapid prototyping machine, are shown in Fig. 1. The white nozzles (back row) are designed for a nozzle pressure ratio (NPR) of six, where NPR is defined as the ratio of nozzle supply pressure to atmospheric pressure. The NPR of 6 was chosen because that corresponds to the highest testing pressure achievable in the existing blow-down facility. The yellow nozzles (front row) are designed for a higher NPR of 50, which more closely matches flight conditions. Each of these model nozzles has undergone testing in the blow-down facility. Our preliminary results were encouraging and of sufficient merit to warrant the award of both Phase II and a concurrent Phase III supplement.

Schlieren images of the air flow exiting the two centered nozzles from Fig. 1 are shown in Figs. 2 and 3. The NPR 6 centered nozzle is operating with a supply pressure near its design value of 90 psia, and the resulting flow, as shown in Fig. 2, is fully expanded. The NPR 50 nozzle is operating at the maximum available pressure of 120 psia, which is significantly lower than its design value of 750 psia. In this case the resulting flow, as shown in Fig. 3, is highly over-expanded, and the usefulness of the test is questionable. The Schlieren images confirm that pressures higher than are available in our cold-flow, blow-down facility are needed for proper testing of nozzles designed for NPRs greater than 6. For our Phase I testing, using NPR 6 nozzles was entirely appropriate. However, for advanced testing in Phase II, higher supply pressures are needed. The supply pressure of over 750 psia that would be needed to test at flight conditions would necessitate a very expensive redesign of the entire blow-down facility. However,

---

<sup>1</sup> "Optimal Thrust Vectoring for an Annular Aerospike Nozzle," a NASA Small Business Technology Transfer (STTR) project in which Cal Poly has teamed up with Rolling Hills Research Corporation of El Segundo, CA. Phase I was funded for \$100,000 in January 2005 and completed in January 2006; Phase II was funded for \$600,000 in June 2006 and will be completed in June 2008; and a concurrent Phase III supplement for \$40,000 was funded in August 2006.



Figure 1. Scale model aerospike nozzles for cold flow, blow-down testing. The white and yellow nozzles are for NPR 6 and NPR 50, respectively. From left to right, the three pairs of nozzles have the spike centered, offset, and gimbaled by  $10^\circ$ .

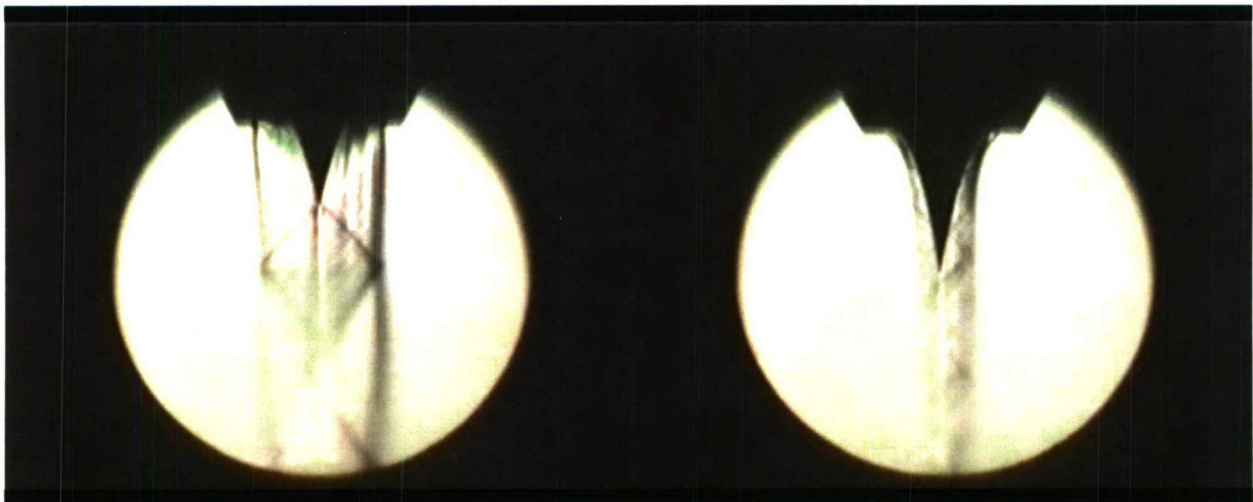


Figure 2. Schlieren for aerospike nozzle with centered spike, NPR 6.

Figure 3. Schlieren for aerospike nozzle with centered spike, NPR 50.

a supply pressure of 300 psia for generating fully developed flow across NPR 20 nozzles is a reasonable and achievable compromise. If our NPR 6 results from Phase I are confirmed in NPR 20 tests in Phase II, full-fledged nozzle development would be appropriate, leading to static hot firings at full pressure and subsequent flight tests for the promising new technology of thrust vectored annular aerospike nozzles.

### *Significance of the C3RP Work Completed To Date*

As was previously mentioned, supply pressures around 300 psia are needed for effective testing of NPR 20 nozzles, which are geometrically much more similar to the NPR 50 flight nozzles than to the NPR 6 nozzles that we can currently test. Because the higher NPR nozzles have a relatively small exit area, the mass flow requirements for the supply system are modest. For these higher NPR nozzles, there are two ways of generating the 300 psia flow: cold flow or hot flow. For cold flow, a manifold fed from a bank of high pressure, compressed air bottles could be added to the existing cold flow, blow-down system. This cold flow approach to testing NPR 20 aerospike nozzles is being carried out in Phase II of our NASA STTR.

A hot flow of high pressure gas to test aerospike nozzles could be generated by using a hybrid rocket motor. Hybrid rocket motors are inherently safe because both the fuel and oxidizer can be stored without special precautions, and they do not burn when brought together without a significant source of heat. The development of a hybrid rocket motor facility for the testing of advanced rocket nozzles is the objective of this C<sup>3</sup>RP project. Not only will this facility be beneficial for the testing of rocket nozzles, but the marriage of hybrid rocket motors and thrust vectored aerospike nozzles is interesting in its own right. The potential for hybrid rocket motors was clearly demonstrated in 2004, when Cal Poly alumni Burt Rutan's SpaceShipOne won the Ansari X-Prize using an HTPB-N<sub>2</sub>O hybrid rocket to reach outer space.

Under this C<sup>3</sup>RP project, the existing Cal Poly hybrid rocket motor facility has been redesigned and upgraded to state-of-the-art research standards. The ultimate propellant combination chosen for this hybrid is liquid nitrous oxide (N<sub>2</sub>O) as the oxidizer and hydroxy-terminated polybutadiene (HTPB) as the fuel, however both HTPB and Plexiglas have been used as the fuel in tests to date. For this hot flow environment, conical rocket nozzles have been fabricated from high-density graphite, however, ultimately aerospike nozzles will be fabricated from high-density graphite, carbon matrix composite materials, superalloys, or combinations of these materials. To accomplish this transformation from our existing small scale hybrid rocket motor facility to a research-grade hybrid rocket motor facility, the following tasks have been performed: 1) the facility has been converted to using liquid N<sub>2</sub>O as the oxidizer and HTPB as the fuel, which allows for significantly higher nozzle pressure ratios and higher thrust; 2) the hybrid rocket motor chamber has been redesigned to accommodate these more powerful burns and the associated higher chamber pressures; 3) the redesigned hybrid rocket motor has been installed on an existing six-axis thrust measurement stand similar to the one used in our cold-flow blow-down tests, and 4) hot fire tests of this hybrid rocket motor have been conducted.

The centerpiece of this project is the design, construction, and operational debugging of the hybrid rocket motor, which is shown schematically in Fig. 4. Design renderings are shown in Figs. 5-6, and photographs of the actual rocket motor are shown in Figs. 7-8. Twenty-two hot fire tests have been conducted to date. The majority of our tests have used Plexiglas as the fuel in the hybrid rocket motor configuration shown in Fig. 5, that is, a conical nozzle and a single hollow cylindrical fuel grain. Through an iterative, learn-as-you-go development process in this testing sequence, we have carefully and conservatively increased the oxidizer flow rate and therefore chamber pressure.

The results from two such hot-fire tests are shown in Fig. 9. The results shown in Fig. 9a are for an early test firing in which stable combustion was achieved yielding a peak chamber pressure of just over 200 psia and a steady-state chamber pressure of approximately 185 psia. The corresponding peak thrust was 58 lbs. Also, the mixture ratio (oxidizer/fuel) was 4.17, as compared to the stoichiometric mixture ratio of 5.28. In this non-oxidizing environment, the high-density graphite nozzle showed no measurable throat

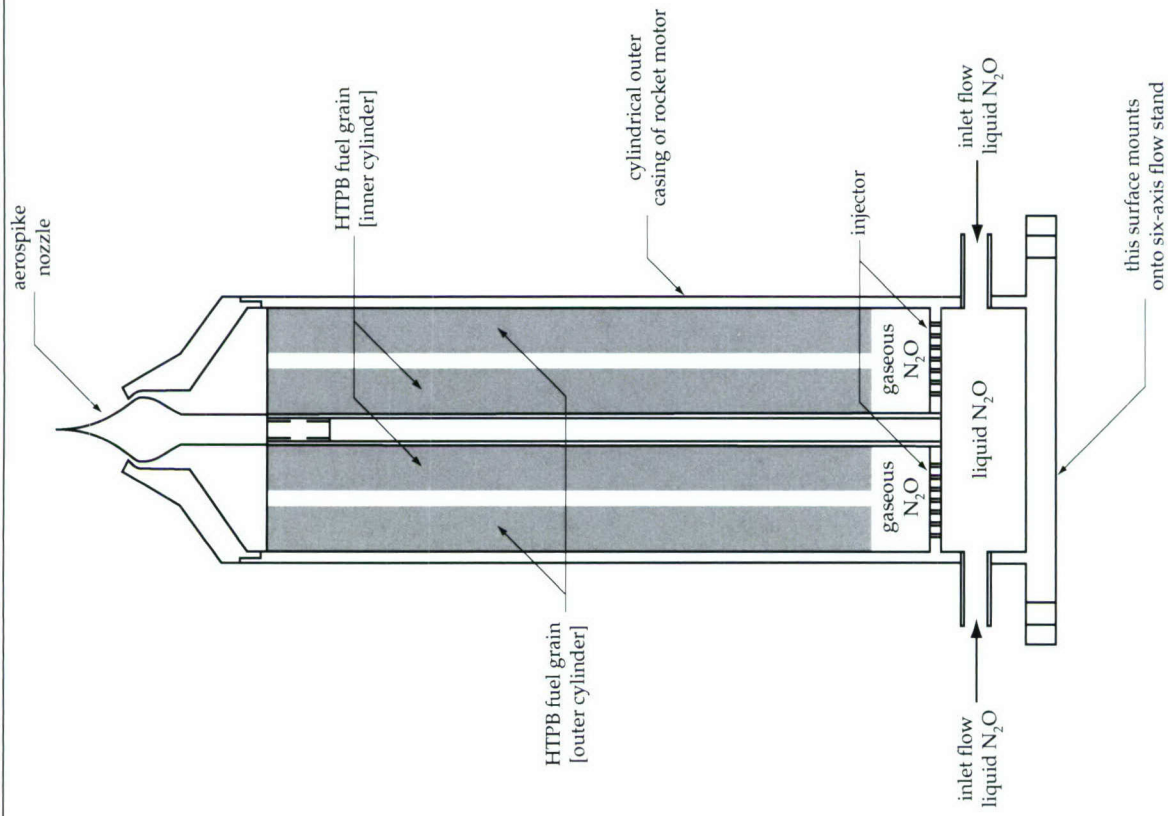


Figure 4: Schematic diagram of the  $N_2O$  - HTPB hybrid rocket motor, shown with an annular aerospike nozzle.

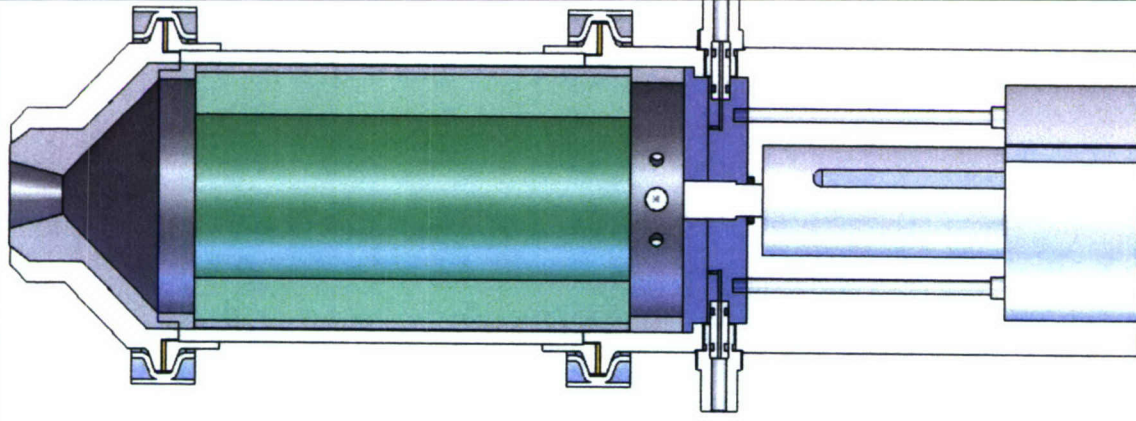


Figure 5. Design rendering for our hybrid rocket motor with a conical nozzle and a single hollow cylindrical fuel grain.

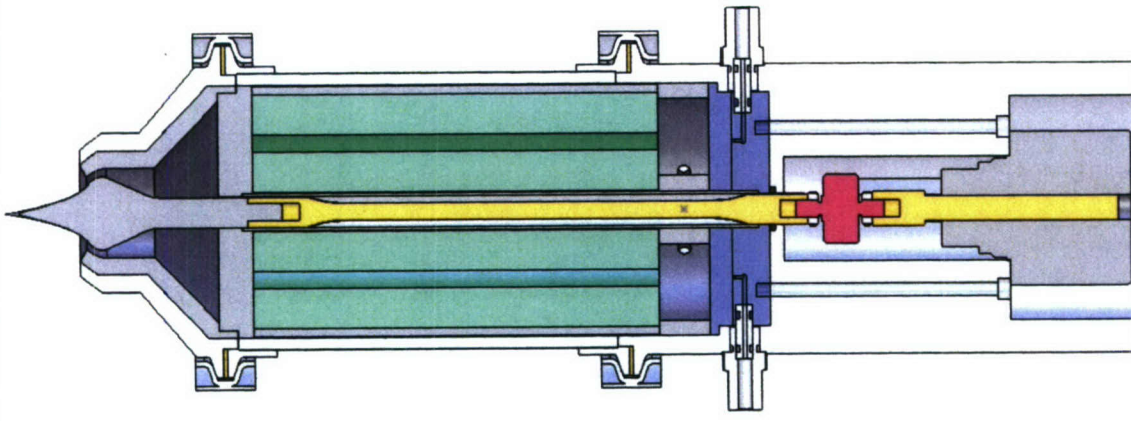


Figure 6. Design rendering for our hybrid rocket motor with an aerospike nozzle and an annular combustion area formed by two fuel grains.

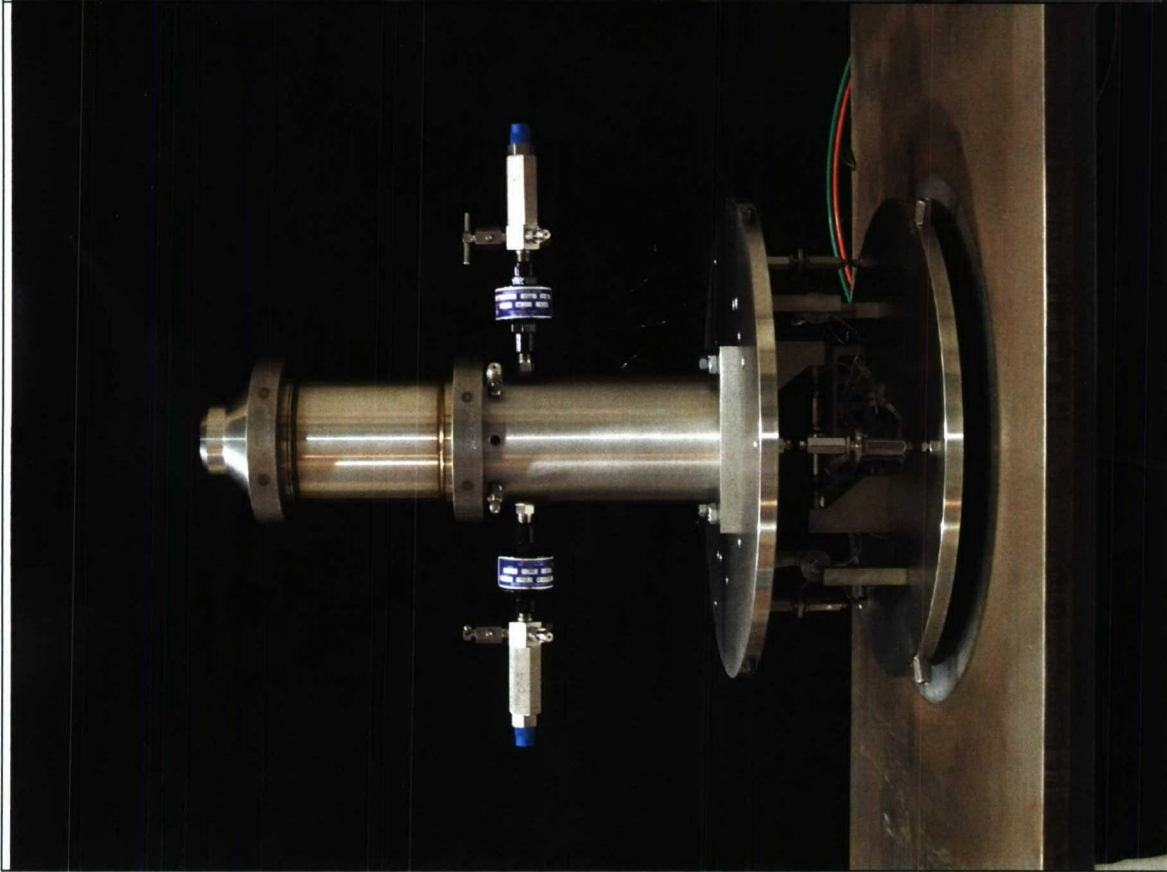


Figure 7. Our hybrid rocket motor mounted on the six-axis thrust stand.

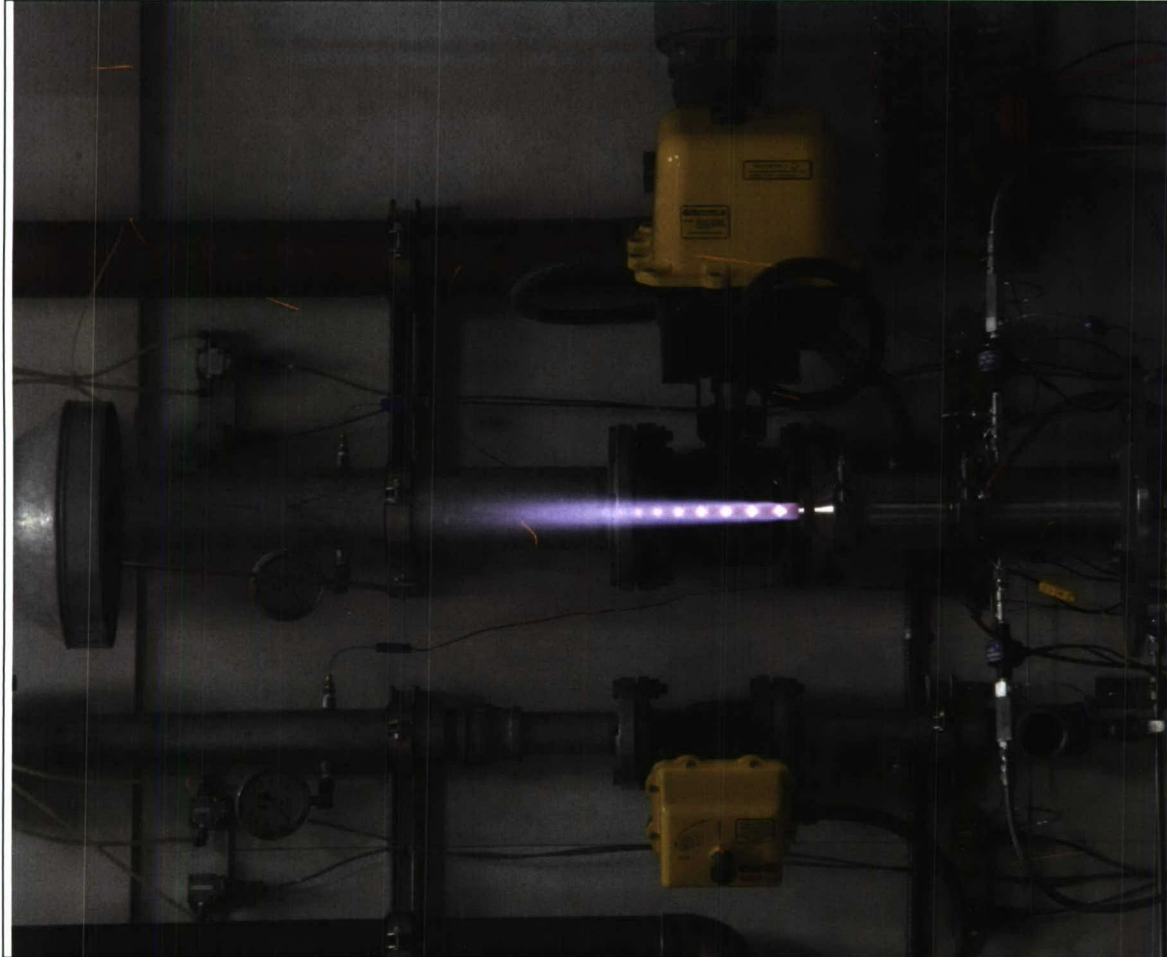


Figure 8. Hot fire testing of our hybrid rocket motor.

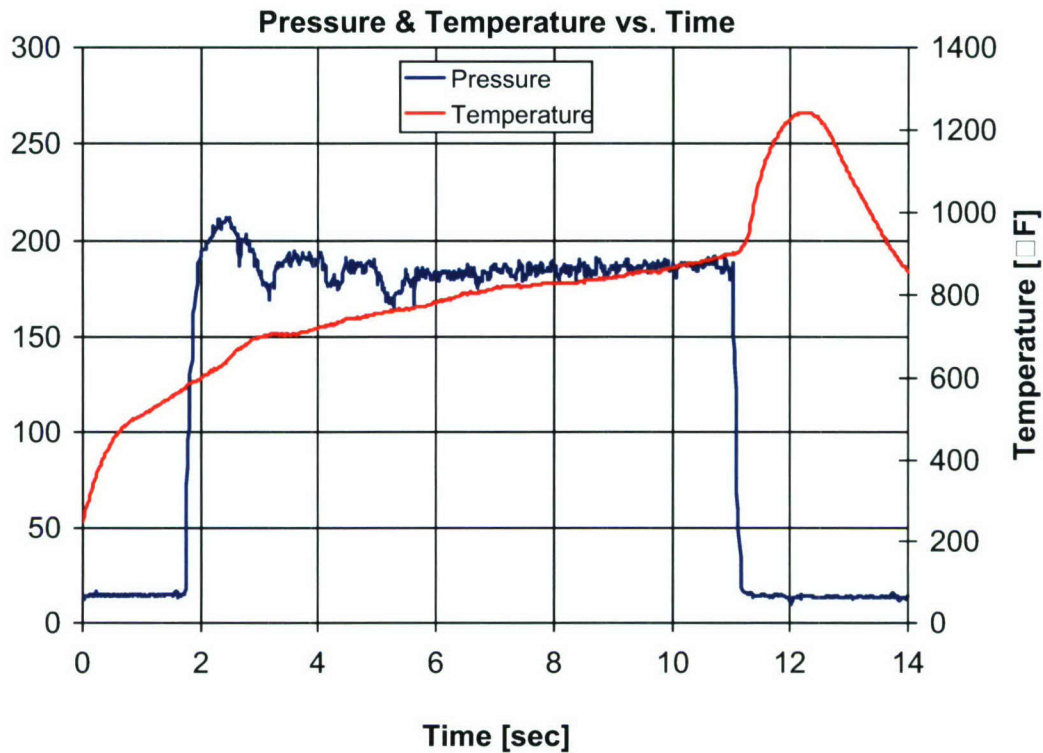


Fig. 9a. Mixture ratio of 4.17, which is slightly below the stoichiometric mixture ratio of 5.28.

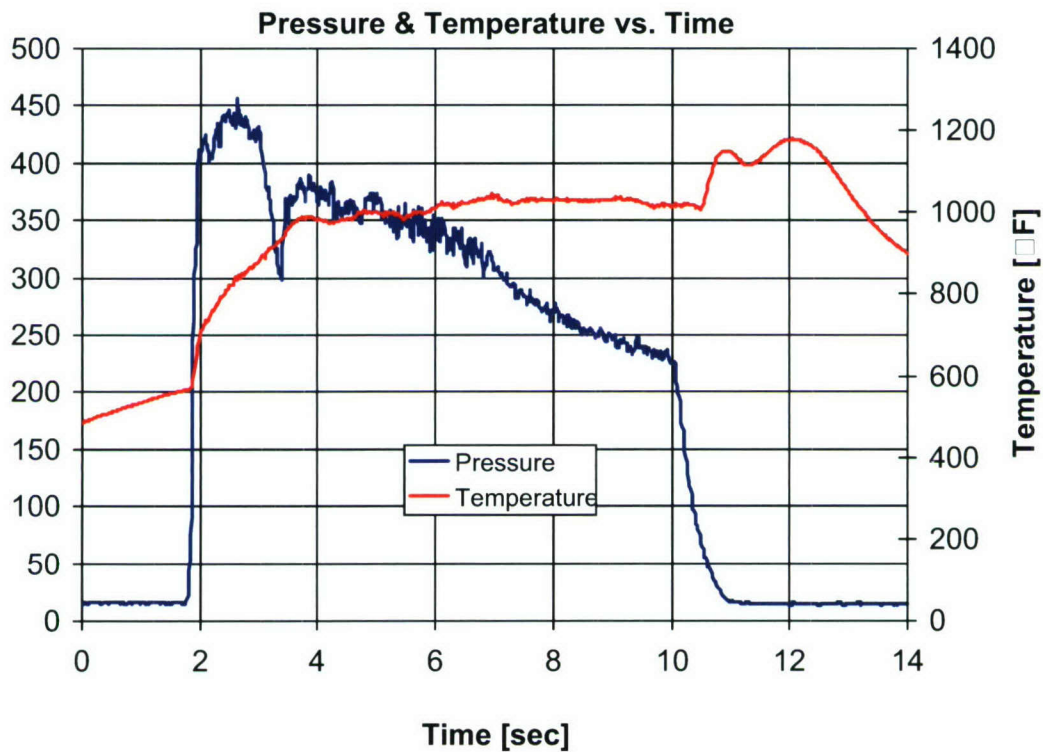


Fig. 9b. Mixture ratio of 8.66, which is significantly above the stoichiometric mixture ratio of 5.28.

Figure 9. Chamber pressure and chamber temperature versus time for hot-fire tests using liquid  $N_2O$  as the oxidizer and Plexiglas as the fuel.

erosion over the course of this nine-second burn. The flatness of the chamber pressure vs. time trace is due to the lack of erosion in the nozzle throat.

The results shown in Fig. 9b are for a recent test firing in which stable combustion was achieved yielding a peak chamber pressure of 450 psia and an average chamber pressure of approximately 325 psia. The corresponding peak thrust was 133 lbs. Also, the mixture ratio (oxidizer/fuel) was 8.66, as compared to the stoichiometric mixture ratio of 5.28. In this highly-oxidizing environment, the high-density graphite nozzle showed significant throat erosion corresponding to a 65% increase in nozzle throat area over the course of this eight-second burn. The continual decrease in chamber pressure vs. time is a consequence of the significant erosion in the nozzle throat.

In other tests, we have successfully test fired a hybrid rocket motor with an annular combustion area formed by an outer hollow cylindrical fuel grain and an inner solid cylindrical fuel grain, and we have successfully mixed, cured, and test fired a hollow cylindrical fuel grain of HTPB.

### ***Further Work***

In the renewal phase of this project, we have budgeted for roughly 100 additional test firings. In this twelve-month sequence of developmental testing, we will: 1) Continue to develop the annular combustion chamber hybrid rocket motor, which will require that we determine the oxidizer flow rate that will yield a 300 psia chamber pressure for this dual grain configuration; 2) If the dual grain configuration does not yield a mixture ratio less than stoichiometric for 300 psia chamber pressures, we will explore additional modifications to the geometry of the fuel grain so that 300 psia chamber pressures can be achieved in steady state with mixture ratios that are below stoichiometric; 3) Continue to work on perfecting our process to produce HTPB fuel grains without introducing air bubbles; 4) Ensure that we can safely and consistently achieve both a steady-state chamber pressure of 300 psia and a mixture ratio slightly less than stoichiometric for the annular combustion area configuration with HTPB as the fuel and liquid  $N_2O$  as the oxidizer. At this point, we will have a hybrid rocket motor research facility that is truly state-of-the-art. If time and budget allow, we will design, build, and hot-fire test an NPR 20 aerospike nozzle on the hybrid rocket motor configuration with an annular combustion area.

### ***Conclusions***

Taken together these improvements will have a dramatic effect on our research, yielding the following benefits: 1) the capability to investigate the performance of annular aerospike nozzles at higher nozzle pressure ratios than currently can be tested in our large cold-flow, blow-down facility, directly benefiting our research into the optimal thrust vectoring and throttling of annular aerospike nozzles; 2) the capability to investigate the performance of our aerospike nozzles (and the mechanisms that move them for thrust vectoring) in a high temperature environment that is reducing rather than oxidizing; 3) the capability to access how well our cold flow nozzle results extrapolate to the corresponding hot flow conditions; and 4) an outstanding chance of success in our upcoming pursuit of new NASA STTR support for thrust vectoring and throttling of annular aerospike nozzles for hybrid rocket motors (proposal due date: September 2007).

## **Effect of reflow profile and thermal shock on lead-free solder joint reliability**

Project Investigator:

Jianbiao Pan  
Industrial & Manufacturing Engineering  
California Polytechnic State University  
San Luis Obispo, California

## **Abstract**

Reflow profile has significant impact on solder joint performance because it influences wetting and microstructure of the solder joint. The purpose of this study is to investigate the effects of reflow profile and thermal shock on lead-free solder joint reliability, specifically shear strength and the intermetallic compound (IMC) thickness.

Test boards were assembled with four different sized surface mount chip resistors (1206, 0805, 0603 and 0402). Nine reflow profiles for SAC 305 and nine reflow profiles for SnPb were developed with three levels of peak temperature (12°C, 22°C, and 32°C above solder liquidus temperature, or 230°C, 240°C, and 250°C for SAC 305; and 195°C, 205°C, and 215°C for SnPb) and three levels of time above solder liquidus temperature (30 sec., 60 sec., and 90 sec.). Half of the test vehicles were then subjected to air-to-air thermal shock conditioning from -40 to 125°C for 500 cycles.

The shear strength data were analyzed using the Analysis of Variance (ANOVA). The fracture surfaces were studied using a Scanning Electron Microscopy (SEM) with Energy Dispersive Spectroscopy (EDS). It was found that thermal shock degraded both SnPb and SnAgCu joints shear strength, and that the effect of thermal shock on solder joint shear strength is much more significant than that of reflow profile. The SnAgCu solder joints have weaker shear strength than the SnPb solders. SnAgCu solder joint after thermal shock retains more of its shear strength than that of SnPb for small components and vice versa for larger components.

IMC thickness was measured using a SEM with EDS as well. The results show that the IMC thickness increases with higher reflow peak temperature and longer time above liquidus together with thermal shock testing.

## **1. Introduction**

The solder joint between the component and the printed circuit board (PCB) should provide not only a reliable electrical interconnection, but also a reliable mechanical interconnection. The transition from through-hole technology to surface mount technology has increased the importance of reliable solder joint shear strength, because the solder joint itself must support a shear force due to mechanical shock and/or thermal stress.

In response to the European Union (EU) Restriction of Hazardous Substances (RoHS) and other countries' lead-free directives, the electronics industry is transiting to lead-free soldering to meet the regulatory requirements to eliminate lead from solders used in consumer electronics products. SnAgCu lead-free solder alloy has been widely adopted as an alternative to eutectic tin-lead (SnPb) solder (Handwerker, 2005; Nurmi, et al., 2005; Puttlitz, 2004). But the reliability of SnAgCu solder interconnection is still not well understood (Shangguan, 2005), especially in high-reliability applications such as military and aerospace devices.

Long-term reliability of solder interconnections depends on the intermetallic compound (IMC) formation (Miric and Grusd, 1998). IMC is necessary for a good solder interconnection and its presence shows a bonding layer between the bulk solder and component termination and/or

board substrates. The IMC in SnAgCu soldering mostly consists of  $\text{Cu}_6\text{Sn}_5$  and  $\text{Ag}_3\text{Sn}$ . Solder joint strength can be affected by both lack of IMC as well as excess IMC within the solder joint. If an excessive amount of IMC was formed, its brittle property may raise solder joint fatigue reliability concern.

It is well known that reflow profile influences wetting and microstructure of the initial solder joint, and thus impact on solder interconnection reliability. During the reflow process, the base metal e.g. copper, dissolves into the molten solder and forms the IMC layer at the interface. The effect of reflow profiles on solder joint IMC formation, wetting characteristics, shear strength performance through thermal ageing, and microstructure characterization have been studied widely (Arra, et al., 2002; Bukhari, et al., 2005; Oliver, et al., 2000; Salam, et al., 2004). Salam, et al. (2004) concluded that the peak temperature and the time above liquidus (TAL) during the reflow process are the most critical parameters impacting solder joint reliability. Oliver, et al. (2000) investigated the effect of thermal aging on the shear strength of lead-free solder joints and concluded that the shear strength decreases as the aging time increases. Bukhari, et al. (2005) evaluated the effects of assembly process parameters and thermal aging. Their results indicate that the effects of thermal aging on solder joint shear strength are much larger than the effects of assembly process variables. Harris and Chaggar (1998) considered that the most influential factor that affects the quantity of IMC is the nature of the base materials, followed by the reflow peak temperature and the TAL.

Since the liquidus temperatures of the SnAgCu lead-free solder alloy is higher than that of eutectic SnPb alloy. The liquidus temperature of  $\text{Sn}_{3.8}\text{Ag}_{0.7}\text{Cu}$  and  $\text{Sn}_{3.0}\text{Ag}_{0.5}\text{Cu}$  is between 217 and 219°C, which is 34 – 36°C higher than eutectic SnPb solder. How significant does the higher reflow temperature contribute to the amount of IMC formation has been indicated by Roubaud and Henshall (2001) who concluded that the higher lead-free solder reflow temperature (250°C) did not lead to a significantly higher thickness of IMC layer between the bulk solder and copper substrate in lead-free assemblies. Arra, et al. (2002) found that the thickness of IMC layer between the solder and the component increased when the peak temperature and the time above liquidus increased. They reported that the IMC thickness between a 1206 component and the SnAgCu solder increased from 2-3 $\mu\text{m}$  to 5-8 $\mu\text{m}$  when the reflow time above liquidus was increased from 30 to 90 seconds.

Aging can cause IMC growth as well. Salam, et al. (2004) presented a study of the reflow profile on IMC thickness and reported that the IMC thickness of SnAgCu solder joints increased from 1-2.5 $\mu\text{m}$  to 3-4.5 $\mu\text{m}$  after ageing at 150°C for 300 hours.

The objective of this study was to investigate the effect of reflow profile and thermal shock on the solder joint shear strength and IMC thickness.

## **2. Experiment**

A five-factor factorial design with mixed levels and three replications was selected in the experiment. The input variables are the peak temperature, the duration of time above solder liquidus temperature (TAL), solder alloy, component size, and thermal shock. The peak temperature and the TAL have three levels each and they are: the peak temperature at 12°C,

22°C, and 32°C above solder liquidus temperatures (or 230°C, 240°C, and 250°C for SAC 305 and 195°C, 205°C, and 215°C for SnPb), and the TAL at 30 seconds, 60 seconds, and 90 seconds. Therefore, there are nine reflow profiles for eutectic SnPb solder and nine for Sn3.0Ag0.5Cu (SAC305) solder. Test boards were assembled with four different sizes of pure tin plated surface mount chip resistors (1206, 0805, 0603 and 0402). Here 1206 means a component with a nominal length of 0.12 inch (3.0 mm) and a nominal width of 0.06 inch (1.5 mm). There were fourteen of each resistor size on each board, or 56 components total per board as shown in Figure 1. Three boards were assembled for each experimental run so a total of 54 boards were assembled (3 peak temperature x 3 TAL x 2 solder alloy x 3 replication). The board finish of the test vehicles was Organic Solderability Preservative (OSP). The experimental matrix is listed in Table 1. A 0.1 mm (4 mil) thick laser-cut electro-polished stencil with 1:1 aperture to pad ratio was used. Both SnPb and SAC305 paste were Type 3 with no-clean flux. The reflow process was done in air.

Each board was cut into two identical pieces. The first half of the board stands for the initial time zero and the components on the half of the board were sheared right after assembly. The other half of the test vehicles were then subjected to air-to-air thermal shock conditioning from -40 to 125°C with 30 minute dwell times (or 1 hour per cycle) for 500 cycles. The components were sheared using a Dage-series 4000 shear tester. For the solder paste, reflow profile development, assembly processes, and shear testing parameters please refer to the previously published paper (Pan, et al., 2006a; 2006b).

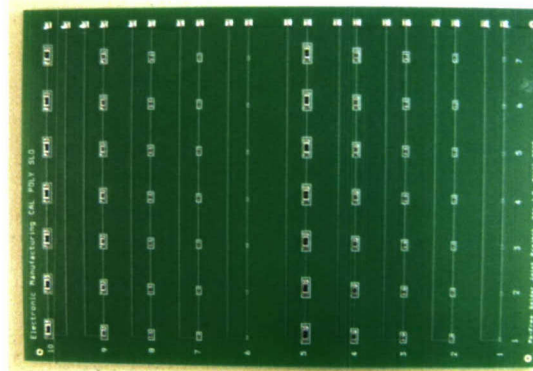


Figure 1. Test Vehicle

Table 1. Experiment matrix

Factors	Levels			
	1	2	3	4
Peak temperature above solder liquidus temperature (°C)	12	22	32	
TAL (sec.)	30	60	90	
Component size	1206	0805	0603	0402
Solder alloy	SnPb	SAC305		
Thermal shock	Before	After		

### 3. Shear Testing Results

The solder joint shear force data were analyzed using analysis of variance (ANOVA). Three assumptions (normality, constant variance, and independence of the residuals) were checked in the ANOVA analysis. The residual plot for shear force (in gram) versus component size shown in Figure 2 indicates that the residual increases slightly as the component size increases. Thus the constant residual variance assumption is not satisfied and a transformation is needed. The square root transformation was found to be appropriate.

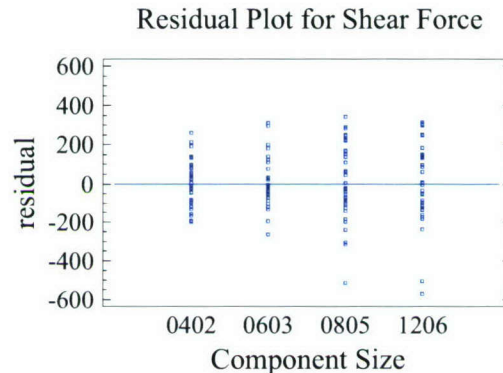


Figure 2. Plot of Residuals vs. Component Size

The ANOVA table for the square root of the solder interconnection shear force is summarized in Table 2. Though all the main factors have a statistically significant effect on the shear force at a 95% confidence level, the F-ratios indicate that the component size, the thermal shock, and the solder alloy are more significant factors compared to the peak temperature and the TAL. The interaction between the component size and the thermal shock is found to be significant as well.

#### ***Effect of Peak Temperature and Time Above Liquidus***

Figure 3 shows that higher peak temperature (32°C above solder liquidus temperature, or 250°C for SnAgCu and 215°C for eutectic SnPb solder alloy) leads to higher shear force. Figure 4 shows that shorter time above liquidus results in higher shear force. As shown in Table 3, higher peak temperature and shorter time above liquidus means a faster cooling rate in the reflow process. The faster cooling rate results in fine microstructure of the solder joint. This may be the reason that higher peak temperature and shorter time above liquidus lead to high shear force. It should be pointed out again that the effects of the peak temperature and the TAL on solder joint shear force is smaller than that of thermal shock.

Table 2. ANOVA for the Square Root of Shear Force

Source	Sum of Square	Df	Mean Square	F-Ratio	P-Value
<b>Main Effects</b>					
A: Peak Temperature	61.7	2	30.9	18.4	0.0000
B: TAL	33.9	2	16.9	10.1	0.0001
C: Component Size	41469	3	13823	8254.3	0.0000
D: Solder Alloy	893.2	1	893.2	533.4	0.0000
E: Thermal Shock	13137	1	13137	7844.7	0.0000
<b>Interactions</b>					
AB	3.8	4	0.9	0.6	0.6884
AC	9.5	6	1.6	1.0	0.4655
AD	19.2	2	9.6	5.7	0.0044
AE	7.2	2	3.6	2.1	0.1223
BC	24.8	6	4.1	2.5	0.0285
BD	21.2	2	10.6	6.3	0.0025
BE	3.6	2	1.8	1.1	0.3410
CD	400.6	3	133.5	79.7	0.0000
CE	3400	3	1133.3	676.8	0.0000
DE	27.9	1	27.9	16.7	0.0001
Residual	172.5	103	1.67		
Total (corrected)	59685	143			

Table 3. Cooling Rate for Different Reflow Profiles

Peak Temperature above the solder liquidus point (°C)	TAL (sec.)	Cooling rate for SnPb (°C/sec)	Cooling rate for SAC305 (°C/sec)
12	30	2.7	3.1
12	60	2.2	2.5
12	90	2.3	2.1
22	30	2.7	3.0
22	60	2.2	2.4
22	90	2.2	2.0
<b>32</b>	<b>30</b>	<b>2.9</b>	<b>3.6</b>
32	60	2.3	2.5
32	90	2.4	2.2

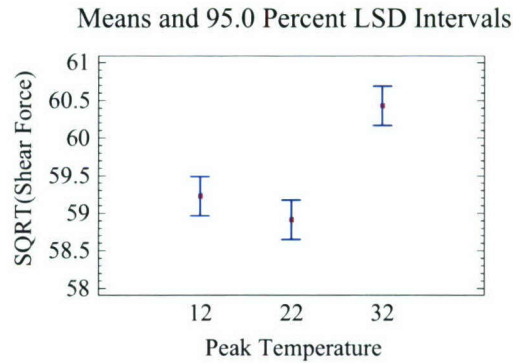


Figure 3. Effect of Peak Temperature on Shear Force

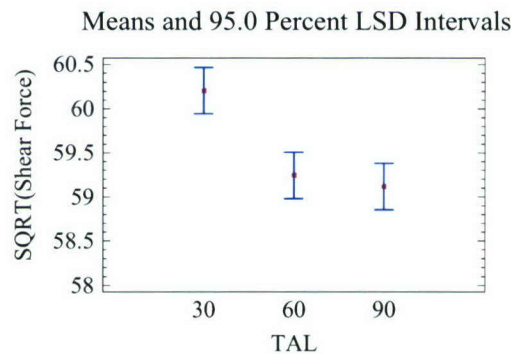


Figure 4. Effect of TAL on Shear Force

### ***Effect of Component Size and Solder Alloy***

Figure 5 shows that the shear force increases with the size of the component. This is expected because the shear force equals to the shear strength of solder alloy times the solder joint wetting area and larger components have a larger solder wetting area.

Figure 6 indicates that the SnAgCu solder joints have weaker shear strength than the SnPb solders. This result is consistent with Oliver, et al. (2000), but different from Sampathkumar, et al. (2005). It should be pointed out that the shear strength of solder alloy depends on the microstructure of the solder joint, which can be influenced by the reflow profile. The published shear strength of eutectic SnPb bulk solder alloy is 45.5MPa and that of Sn3.8Ag0.7Cu alloy is 63.8 MPa (Siewert, et al., 2002). But the shear strength result from bulk solder specimens may not represent the behavior of actual solder joints where the scale of the microstructure may have an effect (Rodgers, et al., 2005). Another possibility could be due to the fact that the wetting of SAC305 was reportedly worse than SnPb (Sattiraju, et al. 2002).

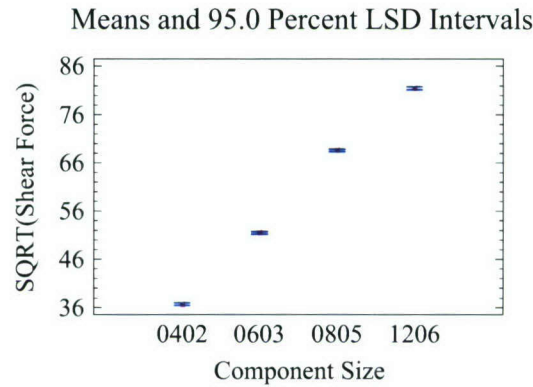


Figure 5. Effect of Component Size on Shear Force

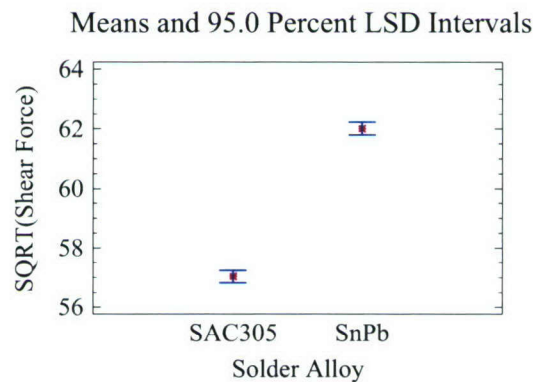


Figure 6. Effect of Solder Alloy on Shear Force

### ***Effect of Thermal Shock***

The interaction plots between the component size and the thermal shock in Figures 7 and 8 show that shear force after thermal shock decreases more as the component size increases. This could be explained due to the larger thermal stress/strain in larger components during the thermal shock testing caused by the thermal expansion mismatch between the component and the board.

It is interesting to note that the shear force loss ratio (the shear force after thermal shock divided by the shear force before thermal shock) for the two solder alloys is different. Figure 9 shows that for small components, SAC305 solder retains more of its strength after heat shock, while for larger components, SnPb solder retains more strength. Table 4 lists the shear forces before thermal shock, after thermal shock, and the shear strength loss ratio for both SAC305 and SnPb solder joints. This result implies that the reliability of SnAgCu alloys outperform SnPb for small components and vice versa for large components. This conclusion is consistent with the thermal fatigue experimental results that suggested that SnAgCu alloys outperform SnPb at low strain applications and vice versa at high-strain amplitude applications (Pan, et al. 2004; 2005; Lau, et al. 2004; Nurmi, et al., 2003; Suhling, et al. 2004; Farooq, et al., 2003).

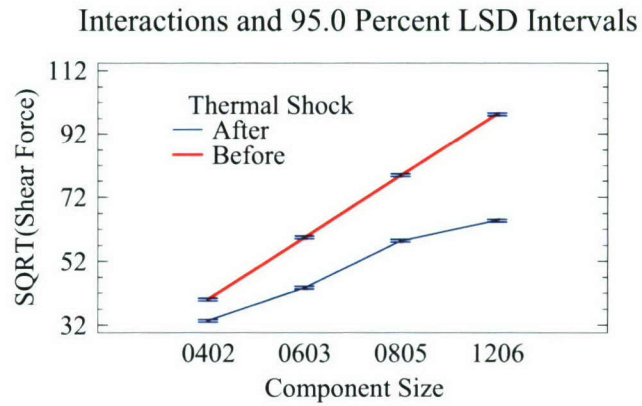


Figure 7. Interaction of Shear Force between Component Size and Thermal Shock

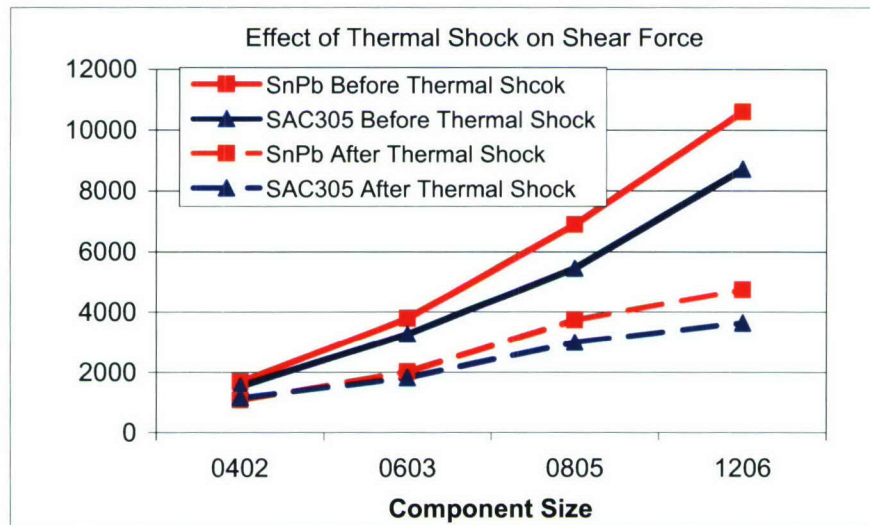


Figure 8. Effect of Thermal Shock on Shear Force

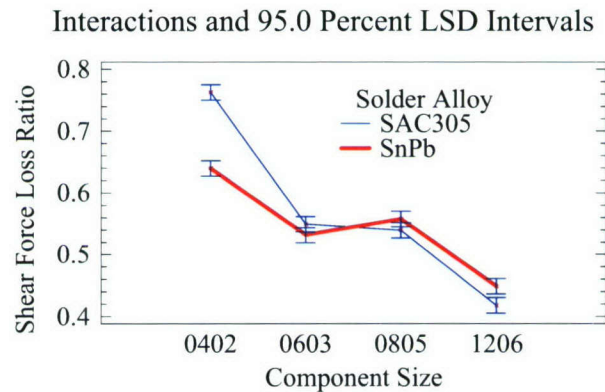


Figure 9. Interaction of Shear Force Loss between Component Size and Solder Alloy

Table 4. Shear Force Loss after Thermal Shock

Paste	Comp. Size	Before Thermal Shock (grams)	After Thermal Shock (grams)	Strength Loss
SAC305	0402	1520	1154	0.759
SAC305	0603	3261	1809	0.555
SAC305	0805	5442	2999	0.551
SAC305	1206	8715	3642	0.418
SnPb	0402	1673	1075	0.643
SnPb	0603	3787	2005	0.529
SnPb	0805	6895	3728	0.541
SnPb	1206	10589	4741	0.448

### ***Fracture Surface SEM/EDS Analysis***

The fracture interfaces of SAC305 solder joints and SnPb solder joints after thermal shock are shown in Figure 10 and Figure 11, respectively. It was observed that the fracture interfaces vary for both SnPb joints and SAC305 joints. In most cases, the fracture occurred partially in the component metallization (Silver layer and or Nickel layer) and partially in the bulk solder joints, which is the same as the fracture surface of solder joints right after assembly (Pan, et al., 2006a; 2006b). In other cases, the fracture occurred in the bulk solder joints only.

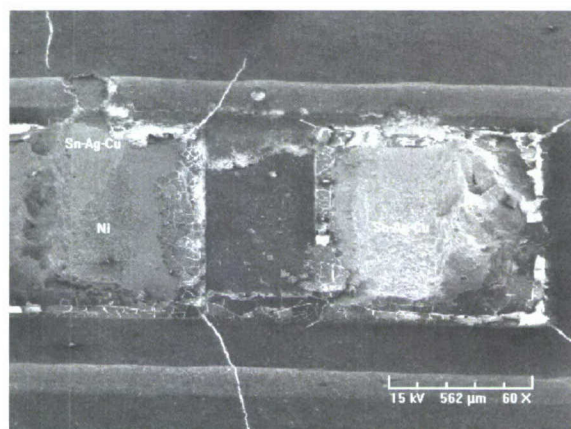


Figure 10. Fracture Surface of SAC305 Solder Joint After Thermal Shock



Figure 11. Fracture Surface of SnPb Solder Joint after Thermal Shock

#### 4. IMC Measurement Results

Samples both from the initial time zero and after thermal shock were cross-sectioned to measure the IMC thickness. The samples were encapsulated in a mixture of epoxy resin and hardener. Care was taken during grinding to not put excessive pressure on the sample to prevent the different metal layers be smeared. The grit size started from 120, following by the number 320, 600, 800, 1200, 2400, and 4000. For each grit size, the technique was to hold the sample in one direction with a scratch pattern opposite to the previous one. The samples were then polished using 0.3 and 0.05 alumina slurries. For the polishing steps, the samples were rotated against the wheel rotation. The last steps were the etching with 50-50 of  $\text{NH}_4\text{OH}$  and  $\text{H}_2\text{O}_2$  then sputter coating with approximately 100 Angstroms of platinum.

The IMC thickness was measured using Scanning Electron Microscopy (SEM) at 5000X magnification with Energy Dispersive Spectroscopy (EDS). There were five measurements on each sample and one sample for each reflow profile. To keep consistency, only the 0603 resistor from each reflow profile was cut out for IMC thickness measurement. Although the IMC formed at both the board pad and component terminal side of the solder joint, only the IMC layer at the board side was measured. This was because the IMC layer on the component side usually was very thin and not easy to distinguish for measurements. Figure 12 shows a sample image of a cross-sectioned sample, and Figure 13 is a magnified image from the rectangular area of Figure 12 showing the solder joint layer structure where the IMC layers can be seen both at the terminal and board pad sides of the joint.

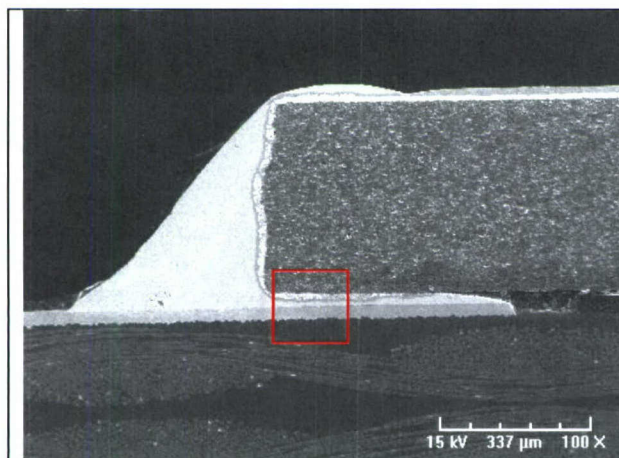


Figure 12. 0603 Chip Cross-section Sample (100X)

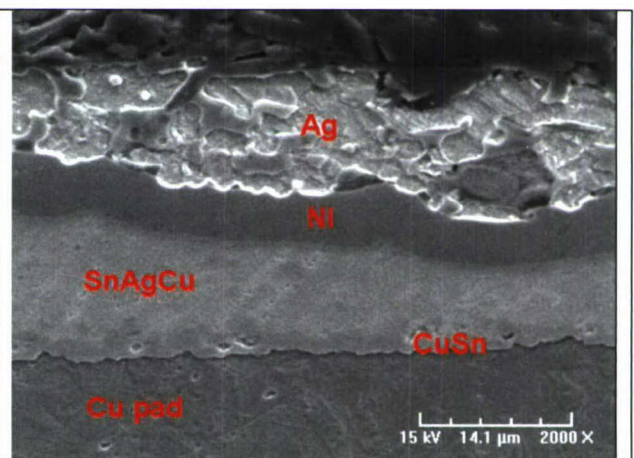


Figure 13 0603 Chip Cross-section Sample (2000X)

#### *IMC Thickness vs. Reflow Profile*

The IMC thicknesses found for this study are summarized in Table 5. Figure 14 shows the SAC305 solder joint IMC thickness before thermal shock increased as the peak temperature and the time above liquidus increased. The IMC thickness increased by 66% when the reflow peak temperature increased from 230°C to 240°C whilst the time above liquidus was kept the same at 90 seconds. But the IMC thickness did not increase when the reflow peak temperature changed from 240°C to 250°C with the same TAL. It was found that there was about a 30% increase in

IMC thickness when the TAL increased from 30 sec. to 90 sec. at the peak temperatures of 240°C and 250°C.

Table 5. IMC Thickness of SAC305 Solder Joints for Different Reflow Profiles

Reflow Profile		IMC Thickness ( $\mu\text{m}$ )			
Peak Temperature ( $^{\circ}\text{C}$ )	Time above Liquidus (sec.)	Before Thermal Shock		After Thermal Shock	
		Mean	Standard deviation	Mean	Standard deviation
230	30	1.12	0.04	1.88	0.12
	90	1.14	0.15	1.83	0.18
240	30	1.45	0.03	1.87	0.54
	90	1.90	0.14	2.47	0.42
250	30	1.43	0.08	1.97	0.38
	90	1.91	0.14	2.17	0.51

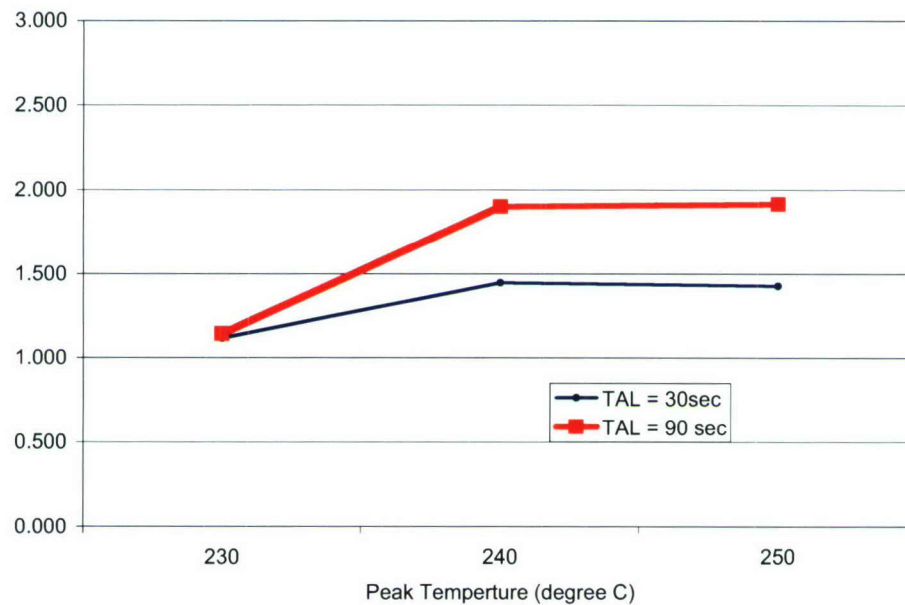


Figure 14. IMC Thickness of SAC305 Solder Joints before Thermal Shock

### ***IMC Thickness vs. Thermal Shock***

IMC thickness comparisons before and after thermal shock from  $-40$  to  $125^{\circ}\text{C}$  for 500 cycles are shown in Figure 15. All the samples appeared a noticeable increase in IMC thickness after thermal shock. There was an over 60% increase in IMC thickness after thermal shock when samples were reflowed at the peak temperature of  $230^{\circ}\text{C}$ . But much smaller increases were found after thermal shock when the initial IMC thickness reached  $2\mu\text{m}$ . This may be explained by the fact that a thick IMC layer would limit further dissolution of Cu into Sn to form the IMC. Figures 16 and 17 are the cross-sectioned images of 0603 components reflowed at the peak temperature of  $240^{\circ}\text{C}$  and the TAL of 30sec., where the IMC thickness of Cu-Sn was measured

by vision gage software. It clearly shows that the IMC thickness after thermal shock was thicker than that before thermal shock.

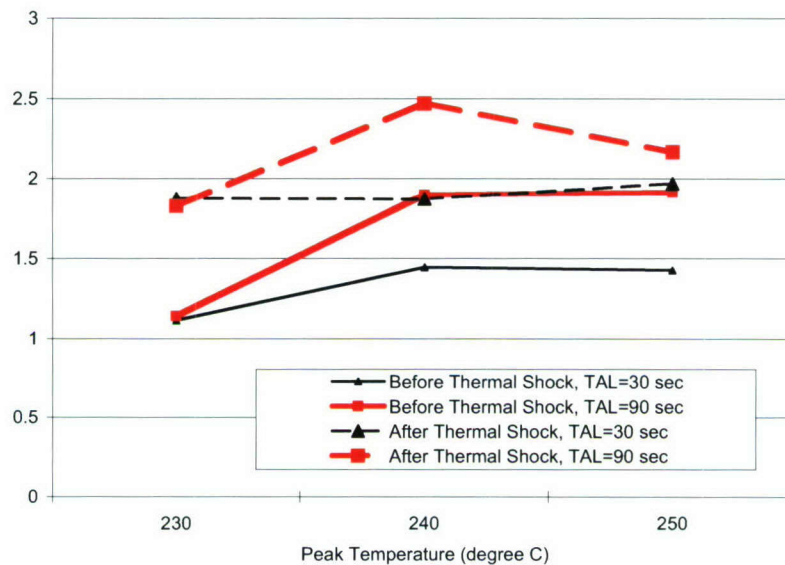


Figure 15. IMC Thickness Comparison Before and After Thermal Shock

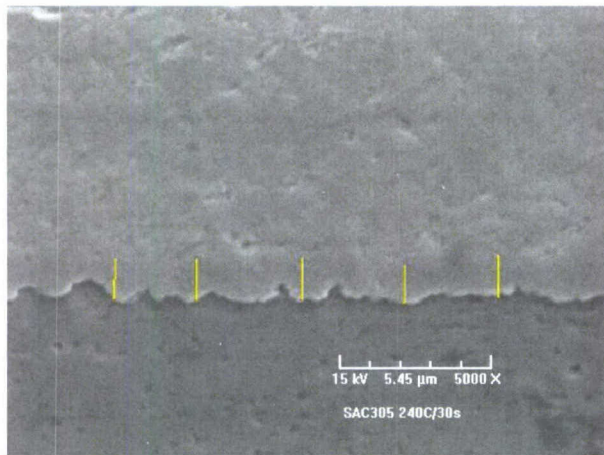


Figure 16. IMC Thickness Before Thermal Shock

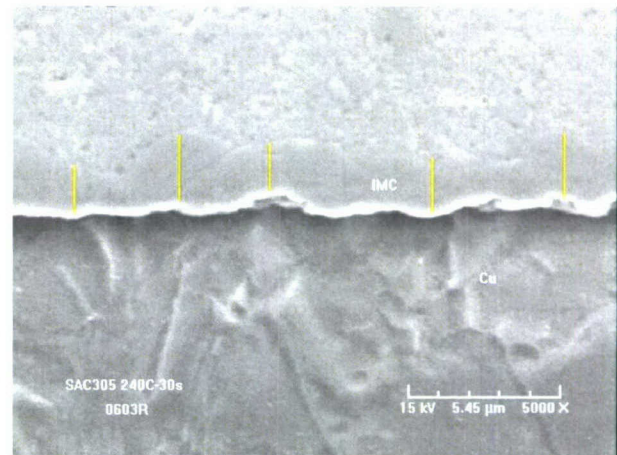


Figure 17. IMC Thickness After Thermal Shock

### ***IMC Thickness vs. Solder Alloy***

The IMC thickness of SnPb and SAC305 solder joints are compared in Table 6. It shows that the IMC of SAC305 solder joints is slightly thicker than that of SnPb solder joints when reflowed at the same peak temperature above liquidus and the same time above liquidus. The results suggest that the high reflow temperature of SnAgCu solder leads to a thicker IMC. It could be explained by the copper dissolving faster at the high reflow temperature and the higher tin content of 96.5Sn3Ag0.5Cu compared with Sn37Pb resulting in more IMC formation.

Table 6. IMC Thickness Comparison between SAC305 Solder Joints and SnPb Solder Joints

Reflow Profile		IMC Thickness ( $\mu\text{m}$ )			
Peak Temperature above liquidus ( $^{\circ}\text{C}$ )	Time above Liquidus (sec.)	SnPb Solder Joints		SAC305 Solder Joints	
		Mean	Standard deviation	Mean	Standard deviation
12	30	0.95	0.10	1.12	0.04
	90	1.23	0.10	1.14	0.15
22	60	1.17	0.10	1.29	0.10
32	30	1.05	0.08	1.43	0.08
	90	1.49	0.11	1.91	0.14

## 5. Summary and Conclusions

The shear strength and IMC thickness of both SnPb and SAC305 solder joints under different reflow profiles were compared. From this study, the following conclusions can be drawn:

- 1) Thermal shock (or loading condition) has significant effect on shear strength degradation of both SnPb and SnAgCu joints.
- 2) The effect of thermal shock on solder joint shear strength is much more significant than that of reflow profile.
- 3) Higher peak temperature and shorter time above liquidus lead to higher shear strength of solder joints.
- 4) The SnAgCu solder joints have weaker shear strength than the SnPb solders.
- 5) SnAgCu solder joint after thermal shock retains more of its shear strength than that of SnPb for small components (at low-strain amplitude applications) and vice versa for larger components (high-strain amplitude applications).
- 6) The IMC thickness of all samples was below  $3\mu\text{m}$  for both SnPb and SAC305 solder joints reflowed at the peak temperature ranging from 12 to  $32^{\circ}\text{C}$  above liquidus temperature and at the time above liquidus ranging from 30 to 90 seconds, before and after thermal shock.
- 7) The IMC thickness increased as the peak temperature and the time above liquidus increased. But the increase of IMC thickness was not linear with the peak temperature and the time above liquidus.
- 8) Thermal shock from  $-40$  to  $125^{\circ}\text{C}$  for 500 cycles leads to slightly thicker IMC layer. It was also found that there were much smaller increases after thermal shock when the initial IMC thickness had reached  $2\mu\text{m}$ . The result may imply that a thick IMC layer would limit further dissolution of Cu into Sn to form more IMC.
- 9) The IMC of SAC305 solder joints was slightly thicker than that of SnPb solder joints for the same peak temperature above liquidus and for the same time above liquidus. The results suggest that the high reflow temperature of SnAgCu solder and an increased percentage of Sn in SnAgCu compared with Sn37Pb lead to a slightly thicker IMC.

## Acknowledgements

The PI would like to acknowledge the research funding by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855. The PI would like to thank Edward

Clements for shear testing, Tzu-Chien Chou for SEM analysis, Charlson Bernal and Roger Jay at Solectron for assistance in SEM analysis, and Jasbir Bath, Dennis Willie of Solectron Corp. for technical support. The PI also wants to acknowledge support from Henkel Technologies, Irvine, CA for providing the solder paste, a stencil, and shear testing equipment. Special thanks to Flextronics International, DEK, Siemens, and Heller Industries for sponsoring the surface mount assembly line at Cal Poly, where the boards were assembled.

## References

- Arra, M.; Shangguan, D.; Ristolainen, E.; and Lepisto, T. (2002), "Effect of Reflow Profile on Wetting and Intermetallic Formation Between Sn/Ag/Cu Solder Components and Printed Circuit Boards," *Soldering and Surface Mount Technology*, Vol. 14, No. 2, pp.18-25.
- Bukhari, B.; Santos, D.L.; Lehman, L.P.; and Cotts, E. (2005), "Continued Evaluation of the Effects of Processing Conditions and Aging Treatments on Shear Strength and Microstructure in Pb-free Surface Mount Assembly," *Proceedings of the SMTA Pan Pacific Microelectronics Symposium*.
- Farooq, M.; Goldmann, L.; Martin, G.; Goldsmith, C.; Bergeron, C. (2003), "Thermo-Mechanical Fatigue Reliability of Pb-Free Ceramic Ball Grid Arrays: Experimental Data and Lifetime Prediction Modeling," *Proceedings of the IEEE/CPMT Electronic Components and Technology Conference*, New Orleans, LA, pp. 827-831.
- Handwerker, C (2005), "Transitioning to Lead-free Assemblies," *Printed Circuit Design and Manufacture*, March 2005, pp. 17-23.
- Harris, P.G. and Chaggar, K.S. (1998), "The Role of Intermetallic Compounds in Lead-free Soldering," *Soldering & Surface Mount Technology*, Vol. 10, No. 3, pp. 38-52.
- Lau, J.H.; Shangguan, D.; Lau, D.; Kung, T.; and Lee, R. (2004), "Thermal-Fatigue Life Prediction Equation for Wafer-Level Chip Scale Package (WLCSP) Lead-Free Solder Joints on Lead-Free Printed Circuit Board (PCB)," *Proceedings of IEEE Electronic Components and Technology Conferences*, Las Vegas, May, pp. 1563-1569.
- Lau, J.H.; Hoo, N.; Horsley, R.; Smetana, J.; Shangguan, D.; Dauksher, D.; Love, D.; Menis, I.; and Sullivan, B. (2004), "Reliability Testing and Data Analysis of Lead-Free Solder Joints for High-density Packages," *Soldering & Surface Mount Technology*, Vol. 16, No. 2, pp. 46-68.
- Miric, A.Z. and Grusd, A. (1998), "Lead-Free Alloys," *Soldering & Surface Mount Technology*, Vol. 10, No. 1, pp. 19-25.
- Nurmi, S.T.; Sundelin, J.J.; Ristolainen, E.O. and Lepisto, T. (2005), "The effect of PCB surface finish on lead-free solder joints," *Soldering & Surface Mount Technology*, Vol. 17, No. 1, pp. 13-23.
- Nurmi, S.T.; Sundelin, J.J.; Ristolainen, E.O. and Lepisto, T. (2003), "The influence of multiple reflow cycles on solder joint voids for lead-free PBGAs," *Soldering & Surface Mount Technology*, Vol. 15, No. 1, pp. 31-38.
- Oliver J.R.; Liu, J. and Lai, Z. (2000), "Effect of Thermal Ageing on the Shear Strength of Lead-free Solder Joints," *Proceedings of the IEEE International Symposium on Advanced Packaging Materials*, pp. 152-156.

Pan, J.; Wang, J.; and Shaddock, D.W. (2004), "Lead-Free Solder Joint Reliability – State of the Art and Perspectives," *Proceedings of 37th International Symposium on Microelectronics*, Long Beach, CA.

Pan, J.; Wang, J.; and Shaddock, D.W. (2005), "Lead-Free Solder Joint Reliability – State of the Art and Perspectives," *IMAPS Journal of Microelectronics and Electronics Packaging*, Vol. 2, No. 1, pp. 72-83.

Pan, J.; Toleno, B.J.; Chou, T.; and Dee W.J. (2006a), "Effect of Reflow Profile on SnPb and SnAgCu Solder Joint Shear Force," *Proceeding of IPC Printed Circuits Expo, APEX and the Designers Summit 2006*, Anaheim, CA, Feb. 8-10.

Pan, J.; Toleno, B.J.; Chou, T.; and Dee W.J. (2006b), "Effect of Reflow Profile on SnPb and SnAgCu Solder Joint Shear Strength," *Soldering and Surface Mount Technology*, Vol. 18, No. 4, 2006, pp. 48-56.

Puttlitz, K.J. (2004), *Handbook of Lead-Free Solder Technology for Microelectronic Assemblies*, Edited by K. J. Puttlitz and K. A. Stalter, Marcel Dekker, Inc., New York, pp. 10.

Rodgers, B.; Flood, B.; Punch, J.; and Waldron, F. (2005), "Determination of the ANAND Viscoplasticity Model Constants for SnAgCu," *Proceedings of IPACK'2005, ASME InterPACK conference*, San Francisco, CA, USA, July 17-22.

Roubaud, P. and Henshall, G. (2001), "Thermal Fatigue Resistance of Pb-Free Second Level Interconnect," *Proceedings of SMTA International*.

Salam, B.; Virseda, C.; Da, H.; Ekere, N.N.; and Durairaj, R. (2004), "Reflow Profile Study of the Sn-Ag-Cu Solder," *Soldering and Surface Mount Technology*, Vol. 16 No. 1, pp.27-34.

Sampathkumar, M.; Rajesnayagham, S; Ramkumar, S.M.; and Anson, S.J. (2005), "Investigation of the Performance of SAC and SACBi Lead-free Solder Alloys with OSP and Immersion Silver PCB Finish," *Proceedings of SMTA International 2005*, Chicago, IL, USA, Sept. 25-29, pp. 568-575.

Sattiraju, S.V.; Dang, B.; Johnson, R.W.; Li, Y.; Smith, J.S.; and Bozack, M.J. (2002), "Wetting Characteristics of Pb-Free Solder Alloys and PWB Finishes," *IEEE Transactions on Electronics Packaging Manufacturing*, Vol. 25, No. 3, pp. 168-184.

Shangguan, D. (2005), *Lead-free Solder Interconnect Reliability*, ASM International.

Siewert, T.; Liu, S.; Smith, D.R.; and Madeni, J.C. (2002), "Database for Solder Properties with Emphasis on New Lead-free Solders," NIST & Colorado School of Mines, Release 4.0, available at [http://www.boulder.nist.gov/div853/lead free/solders.html](http://www.boulder.nist.gov/div853/lead_free/solders.html)

Suhling, J.C.; Gale, H.S.; Johnson, R.W.; Islam, M.N.; Shete, T.; Lall, P.; Bozack, M.J.; Evans, J.L.; Seto, P.; Gupta, P.; and Thompson, J.R. (2004), "Thermal Cycling Reliability of Lead-Free Chip Resistor Solder Joints," *Soldering & Surface Mount Technology*, Vol. 16, No. 2, pp. 77-87.

Webster, J.; Pan, J.; and Toleno, B.J. (2006), "Investigation of the Lead-free Solder Joint Shear Performance," *Proceedings of the 39<sup>th</sup> International Symposium on Microelectronics (IMAPS'2006)*, San Diego, CA, Oct. 8-12.

**Recycled waste materials as additives to improve the performance of  
soil-cement - A laboratory investigation**

Project Investigator:

Ashraf M. Rahim  
Civil and Environmental Engineering  
California Polytechnic State University  
San Luis Obispo, California

**Abstract**

Research was conducted to investigate the performance of a variety of lean soil-cement mixtures enhanced with industrial byproducts. Lower strength mixture designs were chosen following the recent trend as a method of potentially reducing the reflection of shrinkage cracking through asphalt overlays. Criteria based on the 7 and 28-day compressive strengths, were adopted on which each optimal supplemental material percentage was selected (250 psi @ 7 days and 350-400 psi @ 28 days). A series of unconfined compressive strength tests were performed along with splitting tensile strength tests. A threshold was not set for the splitting tensile test results, but this parameter was determined because tensile strength could help to restrain shrinkage and improve a mixture's post-peak strain behavior (pseudoductile response). ASTM D 559-03 (wetting and drying durability) was performed on supplemental material combinations which met the strength criteria with the minimal cement content. A modified version of ASTM C 490-00a was also performed to analyze shrinkage.

Three industrial byproducts were evaluated in this research— Class F fly ash, recycled carpet fiber, and cement kiln dust (CKD). It was found that both fly ash and CKD could partially replace cement in mixtures while still meeting the strength criteria. Recycled carpet fiber was found to be difficult to effectively disperse in soil-cement and it increased the rate of shrinkage. The baseline design with 5% cement content and no supplemental materials experienced the least shrinkage and evaporative loss. CKD required a higher molding moisture content which led to large evaporative losses as the sample air dried during the shrinkage testing. As a result the CKD mixture experienced the largest shrinkage.

**Acknowledgment**

This report includes the results of a study titled, "Recycled Waste Materials as Additives to Improve the Performance of Soil-cement – A Laboratory Investigation" sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855.

The author wishes to thank Dr. Susan Opava for making this support available. The hard work of graduate student Sean Kirby who worked on this project and input from Jennifer Beck in preparing this report are also acknowledged.

## CHAPTER 1 INTRODUCTION

### 1.1 Background

#### 1.1.1 Soil-Cement

The American Concrete Institute defines soil-cement as a mixture of soil and measured amounts of portland cement and water compacted to a high density (ACI 1990). This definition can be further broadened to include other aggregates besides soil. As in concrete, the soil/aggregate particles are bonded together by the cement paste. However, unlike the case of concrete individual particles are not completely covered by the cement paste. An early example of this type of soil stabilization took place in Sarasota Florida in 1915. A mixture of shells, sand, and portland cement was mixed with a plow then compacted to form an early street. To date, more than 100,000 miles of pavement have been installed using soil-cement as a base. Either a hot-mix bituminous wearing surface (asphalt concrete) or a rigid concrete pavement is normally placed over the stabilized base. Soil-cement is used under concrete pavement to prevent the pumping of fine-grained subgrade soils. Under wet conditions and heavy axle loads, water and subgrade soil can be ejected from unstabilized bases creating voids which leave the rigid pavement above unsupported. Other uses for soil-cement have included slope protection for dams and embankment, liners for channels and reservoirs, and mass soil-cement placements for dikes and foundation stabilizations. Soil-cement is now often used to winterize job sites where it provides a stable work platform for construction equipment before concrete placement (ACI 1990).

A closely related stabilization process is roller compacted concrete (RCC) which uses a controlled-graded stone aggregate instead of soil. RCC has been used to construct new gravity dams, arch dams, and for raising and seismically stabilizing concrete dams. Because RCC is stronger and more durable than soil-cement it is used in more demanding projects. Soil-cement is usually covered with a wear surface when used on roadways while RCC is able to be left exposed (Choi and Hansen 2005). For high traffic highways DOTs prefer an aggregate-cement mix instead of soil (George 2002). The construction techniques are similar for all of these stabilization processes. The aggregate (soil, or some combination of coarse and fine aggregates) and cementitious agents are dry mixed either in situ or at a batch plant. The mixture is then spread out at the site, the moisture level is adjusted, and roller compaction takes place. Often several lifts are completed to reach a final thickness which depends on the design parameters. Once properly compacted and given time to cure soil-cement is generally unaffected by moisture.

Several factors need to be considered to determine whether soil-cement is appropriate at a particular location. Much of the economic advantage a soil-cement system can provide comes from savings provided by not purchasing and transporting imported base material. If a local source of high quality aggregate is available this may be the lowest cost and most durable option. However, many roads are needed in remote regions where this type of base material is simply not available. Research has been ongoing in the southern part of the United States to find a method of restabilizing existing soil-cement road bases which have failed due to shrinkage cracking (Gaspard 2000). If existing gradelines and

drainage can be salvaged, the cost of reconstructing these roadways can be kept at a minimum.

A large expansion of water resource projects in the Great Plains and South Central region of the United States was attempted after WWII. The required local supply of riprap was unavailable. Faced with the prospect of potentially enormous transportation costs the Bureau of Reclamation began an intensive soil-cement testing program. By 1961 there was enough research validating the durability and effectiveness of soil-cement to proceed with construction projects. Since then more than 300 major soil-cement slope protection projects have been built in North America including the upstream facings of dams, channel spillways, coastal shorelines, highway and railroad embankments, and embankments for inland reservoirs. It was found that soil-cement designed with a minimum 7-day compressive strength of 750 psi was adequate to withstand the abrasive force of stormwater flows of 25,000 to 45,000ft<sup>3</sup>/s at velocities of up to 20ft/s. The design requirements (strength & durability of the project) and the quality of locally available material largely determine whether the use of soil-cement will be economically competitive. Different soil types will require varying percentages of increasingly expensive cement to meet minimum strength and durability requirements (ACI 1990).

### ***1.1.2 Industrial Byproducts***

Three industrial byproducts were included in this research— Class F fly ash, recycled carpet fiber, and cement kiln dust (CKD). Historically, Class C fly ash has been the preferred supplemental material for concrete because of its self-cementitious properties, but Class F fly ash is produced in larger quantities and finds fewer commercial applications. The source of fly ash is coal-fired electric power plants where fly ash is produced as a combustion byproduct formed as flue gasses cool. According to the American Coal Ash Association, as of 2001, 58% of the coal combustion products were Class F fly ash and 68% of these required disposal for lack of use (Arora and Aydilek 2005). Fly ash (particularly Class F) is available from most local producers at a very modest cost. Transportation costs largely determine pricing.

Approximately 3% of the municipal solid waste stream generated in the United States is composed of carpets and textiles. Carpet byproducts including production waste and recycled post-consumer carpet are estimated to be produced at a rate of 2,000,000 tons/year (Wang et al. 2000). Recycled carpet fiber is available at very low cost from LA Fiber. As a supplemental material, this product was included on the belief that when properly mixed, fiber reinforcement should interlock soil particles enhancing both shear and tensile performance (Gaspard 2002). Elevated tensile strength should help restrain shrinkage and improve a mixture's post-peak strain behavior (pseudoductile response). Several researchers have found that while fiber inclusion does not enhance unconfined compressive strength, it can improve the mixture's tensile strain capacity and toughness (Gaspard 2002). Toughness is a measure of a material's ability to continue to support loading after it has begun yielding. The area under a load-deformation curve represents the ability of a system to absorb deformation energy (Somayaji 2001). For paving applications an increase in toughness will improve the fatigue resistance of the system. The fiber toughening mechanism has been shown to be effective in concrete, but the results for soil-cement especially with lean cement content are less certain (Sobhan and Mashnad 2002).

Cement Kiln Dust (CKD) is a byproduct of the cement production process. CKD is captured by air pollution control equipment installed on the chimneys of cement kilns. According to the EPA in 1990, the cement industry generated an estimated 12.7 million metric tons of CKD at 111 plants in 38 states of which 4 million metric tons were disposed of in piles, quarries, and landfills. The industry disposed of an estimated 3.3 million metric tons in 1995. There are currently 110 Portland cement plants operating in the United States and Puerto Rico with the highest production in California, Texas, Pennsylvania, and Michigan. CKD has cementitious properties which are dependent on the concentration of hydratable oxides such as free (unreacted) lime or free Magnesia (MgO). But its fine powdery nature (0.3 mm max. size) and caustic properties ( $\text{pH} \approx 12$ ) make CKD difficult material to handle in bulk. If moisture is added to control dust some of the cementitious properties are sacrificed (premature hydration) and the material becomes clumpy [Turner-Fairbanks Highway Research Center (TFHRC) 2002a]. It was included in the scope of this research because it is currently underutilized and has shown potential as a partial replacement for cement in mixtures (Udoeye and Hyee 2002).

## **1.2 Objectives**

The objective of this research is to investigate the effect of using recycled waste/by product materials on the engineering properties of cement-treated bases. Several by product materials that will be investigated include; Coal Ash (Fly Ash), Cement Kiln Dust (CKD) and Carpet Waste. A laboratory testing program will be developed starting with the mixture design for each of the aforementioned waste materials added to the soil cement blend. Unconfined compressive strength as a design criterion will be used to select successful mixes. These mixes will be subjected to durability measuring testing program that include wet-dry cycles.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Background

Soil-cement is a term used to describe a soil or aggregate that has been treated with a relatively small proportion of Portland cement (PCA, 2003). The objective of the treatment is to amend undesirable properties of problem soils or substandard materials so they are suitable for use in construction.

Soil-cement has many properties that recommend it as a flexible pavement base course. The perceived problems with cement stabilized base roads have generally stemmed from the tendency for discrete cracks within the base to propagate through the bituminous wearing surface – generally 38 to 152mm (1.5 to 6 in.) thick – giving rise to maintenance concerns (George, 2001). The cement stabilized layer shrinks due to drying, either from loss of moisture and/or “self desiccation” (moisture depletion resulting from cement hydration). It is argued that shrinkage cracking is a natural characteristic of soil-cement, signifying that the cement is producing a hardened base with significant flexural and tensile strength (Norling, 1973). Should the cracks become wider, however, degradation of the pavement along the cracks not only leads to a rough riding surface but also to delamination of the layers and local failure. The latter phenomenon is reinforced in recent studies (Little et al, 1995 and Kota et al., 1995). For instance, Little, et al. (1995) investigated the performance of several heavily stabilized bases, and concluded that the performance of the sections is dictated by the amount of shrinkage cracking. Wide shrinkage cracks have been singled out as a factor for premature degradation of soil-cement pavements. The wider the crack the more the water infiltration and consequent pumping of the underlying material. With the load-induced stresses increased along the crack edge, secondary cracks begin to appear, typically in the longitudinal direction along the wheel path.

Several procedures/techniques have been proposed for minimizing shrinkage cracks and resulting reflection cracking. For the purpose of this overview, they are grouped into four categories: First, controlling maximum shrinkage and consequent cracking by proportioning materials. Examples include use of minimum content consistent with long-term durability, specifying limiting fines content in the soil, and assuring practical minimum moisture during compaction, to name the important ones. Second, expansive cement, fly ash cement, or secondary additives such as fly ash and a host of other organic compounds have been proposed, again to reduce drying shrinkage (George, 1970). Quality construction, including maximum density close to proctor, proper curing and improving uniformity of mix fall under the third category. Quality construction, including maximum density close to proctor, proper curing and improving uniformity of mix fall under the third category. The fourth category addresses the issues directly by controlling shrinkage cracking in the stabilized layer so that reflection cracking is minimized. The latest category includes: Precracking the cement-stabilized base by delaying placement of surface; Precracking (mechanically) by immediately opening the base to traffic; Controlling cracking by precutting (Colombier and Marchand, 1993); Use of interlayers (surface treatment or stress relieving layers) inhibiting propagation of cracks from the base layer; Use of thicker asphalt concrete (AC) surface; Use of thicker base slab with reduced cement content, and; Prescribing material/methods (such as precracking) that promote numerous minute cracks (microcracks) in contrast to a few wide cracks.

One of the earliest laboratory studies of soil-cement by Davidson et al. (1958) reported that fly ash addition increased the strength of the composite mixture by 28 percent. Fly ash has been used in cement-treated soils with encouraging strength results as well as satisfactory field performance (Kasibati and Conklin, 1994). The Electric Power Research Institute cited various projects in which fly ash was used as an admixture with cement and lime (EPRIPA, 1986). None of those studies attempted to evaluate the effectiveness of fly ash in reducing drying shrinkage and consequent cracking of base layers.

In this proposed research three different recycled byproducts, namely carpet waste, cement kiln dust and fly ash will be used in an attempt to improve the performance of soil-cement.

## CHAPTER 3 EXPERIMENTAL WORK AND RESULTS

### 3.1 Introduction

Laboratory testing was performed with soil obtained from a construction site located on the Cal Poly campus. After preparing the soil by mixing, air drying, and breaking up any clumps with a tamper, the homogenized material was passed through a #4 sieve. An ASTM C 136 standard sieve analysis was performed. This was followed by soil classification according to ASTM D 2487-93 (Unified Soil Classification System). It was necessary to determine the Liquid Limit, the Plastic Limit, and the Plasticity Index of the soil for the classification of the fines portion of the soil. These tests were performed in accordance with AASHTO T89-02 and T90-00. Once the soil type was classified, the optimal moisture content (OMC) required for compaction of the soil-cement mixture was determined using the procedure outlined in ASTM D 698 which describes the proctor moisture-density test.

Soil-cement cylinders were then formed in a 4 inch diameter by 4.6 inch mold compacted in 3 layers at OMC using standard compactive effort. First a set of variable cement content cylinders (4-6% by weight) were created and then tested to find their unconfined compressive strength. This test was performed in general accordance with ASTM D 1633-00, a compressive strength test for molded soil-cement cylinders. Specimens from the same initial set of cylinders were also tested for their split tensile strength in accordance with ASTM C 496-04 designed for concrete. The combination of these two strength test results served as the base values for evaluating the performance of mixtures altered with supplemental materials. Strengths after both 7 day and 28 day curing periods were investigated. The industrial byproducts evaluated included Class F fly ash, recycled carpet fiber, and cement kiln dust (CKD).

Criteria based on the 7 and 28-day compressive strengths were adopted on which the optimal supplemental material percentages were selected. ASTM D 559-03—*Standard Test Methods for Wetting and Drying Compacted Soil-Cement Mixtures* was performed to determine the durability of the different mixtures. A modified version of ASTM C 490-00—*Standard Practice for Use of Apparatus for the Determination of Length Change of Hardened Cement Paste, Mortar, and Concrete* was also performed to analyze the shrinkage experienced by each mixture combination. Table 3-1 details the nomenclature and composition of each mixture evaluated in this research program.

Table 3-1: Mixture nomenclature and composition of all samples.

Mixture Name	Soil, g	Cement (C), g	Fly Ash (FA), g	Carpet Fiber (CF), g	Cement Kiln Dust (CKD), g	Water, g
4C / --	1898	77				169
4.5C / --	1888	86				169
<b>5C / --</b>	<b>1878</b>	<b>96</b>				<b>169</b>
5.5C / --	1869	105				169
6C / --	1859	115				169
4C / 3FA	1839	77	57			169
<b>3C / 6FA</b>	<b>1799</b>	<b>57</b>	<b>115</b>			<b>169</b>
2C / 9FA	1760	38	172			169
5C / 3CF	1878	96		4.7		169
<b>4C / 3CF</b>	<b>1898</b>	<b>77</b>		<b>4.7</b>		<b>169</b>
4C / 3NF	1898	77		6.1		169
3C / 6CF	1918	57		9.4		169
2C / 9CF	1938	38		14.1		169
3C / 6CKD	1799	57			115	194
3C / 9CKD	1740	57			172	207
2C / 6CKD	1819	38			115	194
<b>1C / 9CKD</b>	<b>1780</b>	<b>19</b>			<b>172</b>	<b>207</b>
0C / 12CKD	1740	0			230	220

### 3.2 Soil Classification and Moisture-density

After preparing the soil by mixing, air drying, and breaking up any clumps with a tamper, the homogenized material was passed through a #4 sieve. An ASTM C 136-05 standard sieve analysis was performed and the results are shown in Figure 3-1.

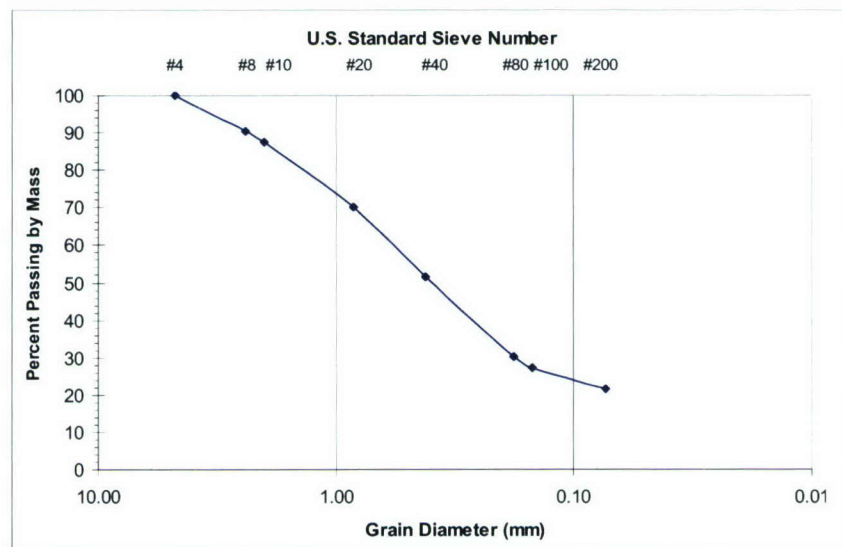
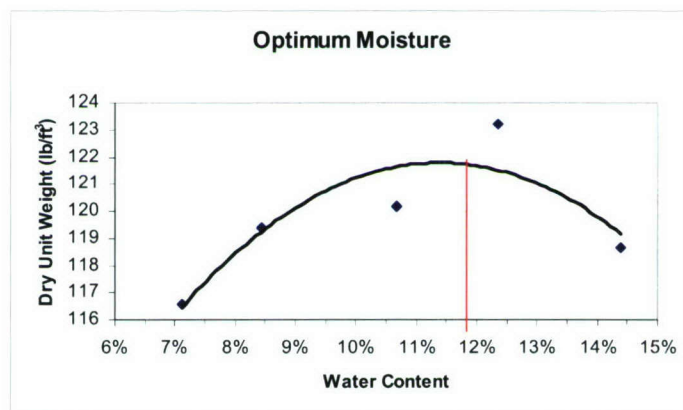


Figure 3-1: Test soil grain size distribution curve.

It was also necessary to determine the Liquid Limit (LL), the Plastic Limit (PL), and the Plasticity Index (PI) of the soil's fines. This testing was performed in accordance with AASHTO T89-02 and T90-00. With the sieve analysis and the LL, PL, and PI the soil was classified according to ASTM D2487—*Classification of Soils for Engineering Purposes (Unified Soil Classification System)*. The soil type was found to be a light brown silty sand (SM) with 21.4% non-plastic fines. The soil's properties are summarized in Table 3-2. (The AASHTO classification was found to be A-2-4). The optimal moisture content (OMC) required for compaction of the soil-cement mixture was determined using the procedure outlined in ASTM D 698—*Test Method for Laboratory Compaction Characteristics of Soil Using Standard Effort* which describes the proctor moisture-density test. For a cement content of 5%, an OMC of about 12% produced a peak dry density of 122.5 lb/ft<sup>3</sup> (Figure 3.2).

**Table 3-2: Soil Properties.**

% Gravel (>#4):	0.0%
% Sand (<#4 and >#200):	78.6%
% Fines (<#200):	21.4%
D <sub>50</sub> :	0.40mm
D <sub>60</sub> :	0.58mm
LL:	18
PL:	Non plastic
PI:	n/a
<b>SM</b>	<b>(Silty Sand)</b>



**Figure 3-2: Moisture-density relationship with standard compactive effort.**

### 3.3 Unconfined Compressive Strength

A set of variable cement content cylinders (4-6% by weight) were molded and tested to find their unconfined compressive strength. Soil-cement cylinders were formed in a 4.58 inch high by 4 inch diameter mold compacted in 3 layers at OMC using standard compactive effort. The specimen's proportions were not the ideal height-to-diameter ratio (h/d) of 2.00 used for Method B of ASTM 1633-00— *Standard Test Methods for*

*Compressive Strength of Molded Soil-Cement Cylinders.* When the h/d ratio is 2.0 the compressive strength values are more accurate because complex stress conditions present at lower h/d values are reduced. This is, however, the standard mold size used for both the moisture-density and the wetting and drying durability tests. Test Method A of ASTM 1633 uses this common mold size which has a h/d ratio of 1.15. For comparison with samples molded with the ideal shape, a correction factor based on ASTM C 42 for concrete would use a strength reduction of 1/(1.1) (ACI 1990). The specimens were not soaked for 4 hours prior to testing which was a deviation from the standard's procedure. The compressive strengths of samples made with differing proportions of supplemental materials and cured for 7 and 28 days are shown in Table 3-3. Mixtures were considered optimal when their 7-day strength exceeded 250 psi and 28-day values were between 350-400 psi with the minimum cement content.

### 3.4 Splitting Tensile Strength

The tensile strengths of selected mixture designs were evaluated using ASTM C 496-04—*Standard Test Method for Splitting Tensile Strength of Cylindrical Concrete Specimens*, a test designed for concrete which measures the tensile performance of a specimen indirectly. In the test setup shown in Figure 3-3 the soil-cement cylinder is held on its side sandwiched in a jig. This produces loading which induces tensile stresses in the plane of applied loading. The test machine used for this procedure was also used for the unconfined compressive strength testing. Plywood bearing plates were added to distribute the load and prevent specimen movement. The highest strength value recorded was the 5% cement sample and neither supplemental material produced any appreciable improvement in performance (Table 3-4).

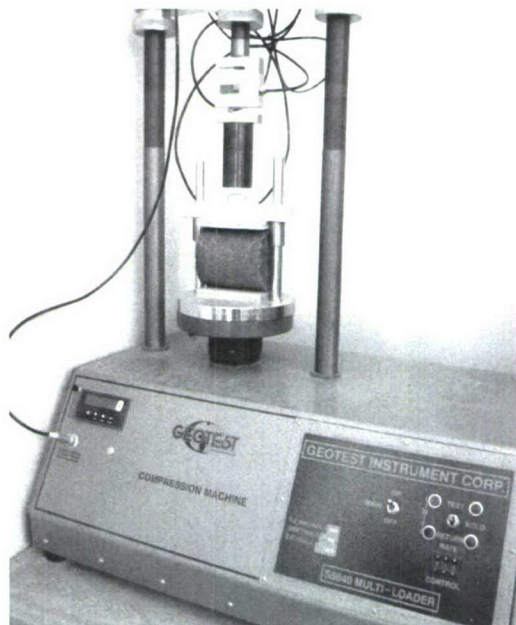
**Table 3-3: Unconfined compressive strength results (bold are considered optimal).**

<b>Mixture</b>	<b>Average 7 day (psi)</b>	<b>Average 28 day (psi)</b>
4C / --	257	342
4.5C / --	303	
<b>5C / --</b>	<b>376</b>	<b>410</b>
5.5C / --	370	
6C / --	394	574
4C / 3FA	373	423
<b>3C / 6FA</b>	<b>266</b>	<b>394</b>
2C / 9FA	201	266
5C / 3CF		447
<b>4C / 3CF</b>	<b>256</b>	<b>360</b>
3C / 6CF	175	285
2C / 6CF	105	187
3C / 6CKD	395	
3C / 9CKD	442	
2C / 6CKD	334	

<b>1C / 9CKD</b>	<b>257</b>	<b>438</b>
0C / 12CKD	221	
Target:	250+	350-400

**Table 3-4: Splitting tensile strength results.**

<b>Mixture</b>	<b>Average 7 day, psi</b>	<b>Average 28 day, psi</b>
4C / --		
4.5C / --		
<b>5C / --</b>	55	111
5.5C / --		
6C / --		
4C / 3FA		90
<b>3C / 6FA</b>		82
2C / 9FA		46
5C / 3CF		
<b>4C / 3CF</b>	51	87
3C / 6CF	22	51
2C / 6CF	15	31



**Figure 3-4: Splitting tensile and unconfined compressive strength test apparatus.**

### 3.5 Wetting and Drying Durability

Specimens molded with the optimal supplemental material percentage were subjected to a durability test to evaluate their performance when exposed to repeated moisture extremes. The goal is to determine if these optimal mixtures can permanently hold a mass together with volume stability. ASTM 559-03—*Standard Test Methods for Wetting and Drying Compacted Soil-Cement Mixtures* was performed to evaluate the durability of the different mixtures. This test consists of subjecting pairs of specimens to a series of 12 cycles of alternating immersion in room temperature water for 5 hours followed by drying in an oven at 160°F for 42 hours. Specimen #1 is the control which is merely measured for mass loss and volume changes. The second specimen, #2 is also measured for mass loss after being exposed to a specified number of heavy brush strokes applied to all surfaces. The standard calls for 2 strokes with a force of 3lb to be applied everywhere using a specified wire brush (Figure 3-4). Different agencies and associations specify levels of acceptable weight loss for samples at the end of the durability tests for different soil types. There are also criteria for the maximum volume and moisture change allowable. The maximum allowable mass loss is 14% for a soil type of SM (silty sand)

according to the PCA (ACI 1990). Volume changes are limited to 2% of the initial volume, and the maximum moisture content cannot exceed the quantity required for saturation at the time of molding PCA (ACI 1990).

**Table 3-5: Durability test results.**

Mixture	Max $\square_d$ (lb/ft <sup>3</sup> )	$\Delta v_{\max}$	$\Delta \text{Mass}_{\max}$	MC <sub>Max</sub>
5C / --	120.5	1.62%	28.8%	9.22%
4C / 3FA	119.1	1.47%	31.3%	9.60%
3C / 6FA	122.2	0.93%	31.5%	9.10%
4C / 3CF	116.4	1.45%	27.9%	11.02%
1C / 9CKD	120.0	1.65%	25.2%	9.75%

maximum volume change-specimen

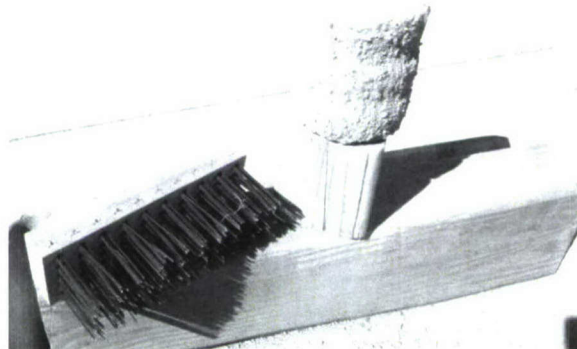
$\Delta v_{\max}$ : 1 (unbrushed)

maximum mass loss-specimen 2

$\Delta \text{Mass}_{\max}$ : (brushed)

maximum moisture content specimen

MC<sub>Max</sub>: 1 (unbrushed)



**Figure 3-4: Wire brush used to abrade surfaces of soil-cement cylinders.**

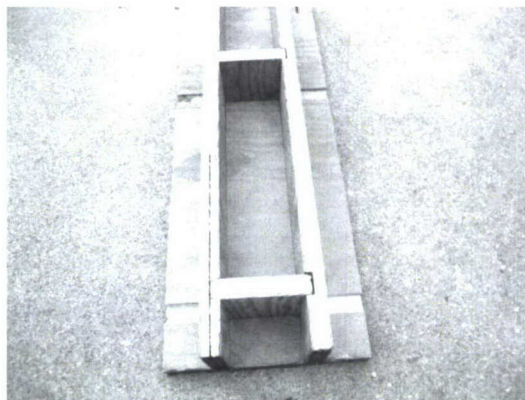
### 3.6 Shrinkage

A modified version of ASTM 490-00a—*Standard Practice for Use of Apparatus for the Determination of Length Change of Hardened Cement Paste, Mortar, and Concrete* was used to measure the shrinkage experienced by each optimal mixture combination. The forming process of the specimens is obviously different between a cast material such as concrete and soil-cement formed by compacting 3 layers of premixed material. The formwork for both types of samples would be the same— a wooden mold (3x3x11 ¼ inches) shown in Figure 3-5 without support clamps attached. Metal studs are installed at the ends of the prism to provide fixed measuring points. These specially manufactured parts act as the interface between the specimens and the test apparatus. Normally they would protrude into the concrete formwork and be cast-in-place. However, because of the necessity for compaction, installing gage studs in soil-cement required more elaborate

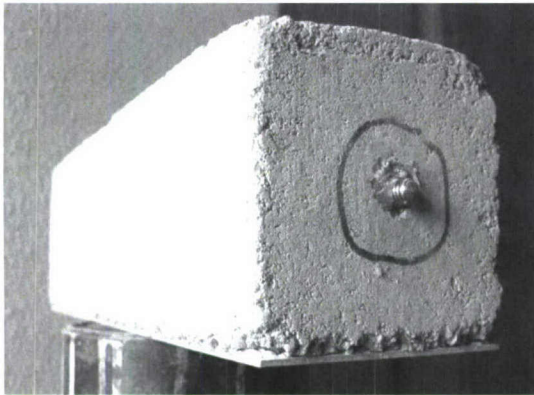
methods. The samples were removed from the molds after curing for 24 hours and 3/8 diameter holes were drilled at each end. The gage studs were then set in these holes and embedded in epoxy. After the initial 24 hour curing when the specimens were removed from the molds to install the gage studs, plastic wrap held in place with masking tape was placed on all exposed surfaces. The wrap was then cut to remove the sheeting from the top surface to simulate the one-sided exposure an installed road bed base would experience.

A sheet metal plate with coarse aggregate was installed at the bottom of the form prior to casting each beam. It served two purposes. The testing was meant to model the type of restrained shrinking present in a roadbed base where the subbase would resist movement from the base. It also acted as a support for the relatively long narrow beams. Both the sheet metal base plate and gage stud are visible in Figure 3-6 and Figure 3-7.

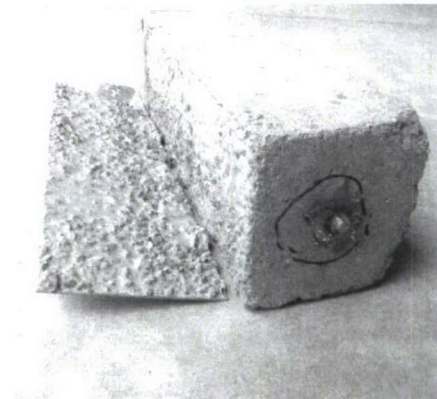
In general a good bond was achieved between the specimen's base and these plates. It was however necessary to epoxy one end of a beam to its plate after the bond was lost. The repeated vertical positioning of specimens in the test apparatus subjected this composite system to a lot of induced stress as the beams were moved around. Initially multiple measurements of the mass and change in length of each specimen were recorded each day while there was a high rate of initial shrinkage. As the changes in length became more gradual sampling was reduced to two measurements per day for the remainder of the experiment's duration which lasted just over two weeks. The specimens were allowed to air dry in a laboratory with temperature and humidity fluctuations. An experimental coefficient of thermal expansion was determined and this was used along with room temperature readings to correct the recorded length measurements. A decreasing ambient temperature will artificially shrink the samples, and the correction will add back this thermally induced loss in length. An increased room temperature has the opposite affect, expanding samples which reduces the apparent shrinkage. The correction will make the sample's length more negative indicating the larger actual shrinkage. A summary of the shrinkage data is presented in (Table 3-6).



**Figure 3-5: Plywood mold for shrinkage samples.**



**Figure 3-6: Stainless steel gage stud protruding from sample.**



**Figure 3-7: Sheet metal base plate with aggregate embedded in epoxy.**

**Table 3-6: Summary of shrinkage data for tested specimens.**

Mixture	Moisture Content (%)	Initial $\mu\epsilon$ /day ((in/in)/day)	Dry Unit Weight $\square_d$ (lb/ft <sup>3</sup> )	Max. Strain $\mu\epsilon$ (in/in)	Maximum Mass Loss (%)
5C / --	12.5	56	126	1313	3.20
3C / 6FA	12.6	211	121.5	2908	4.92
4C / 3CF	12.9	180	121	2669	4.56
1C / 9CKD	15.5	292	117	3874	6.00
4C / --	12.8	116	122	1781	3.71

The test apparatus shown in Figure 3-8 consists of a rigid frame with an accurate ( $\pm .0001$ in.) dial indicator. A reference bar made of a lower thermal expansion alloy with a length of  $11 \frac{5}{8}$  inches was used to recalibrate the comparator at the start of each test. This eliminates any variation in frame's length due to temperature changes in the laboratory.

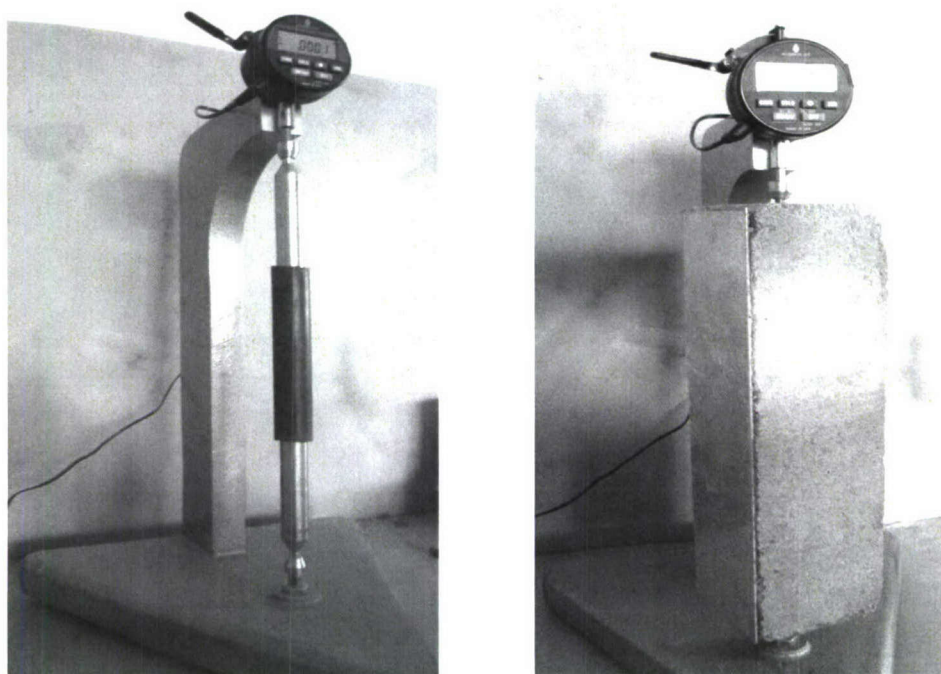


Figure 3-8: Comparator shown with reference bar (left) and test specimen (right)

After the comparator was calibrated the specimen was placed in the frame and rotated until the lowest reading was displayed on the dial indicator. At this position the shortest beam length could be read from on the dial indicator. This would be the change in length from the reference rod experienced by the dial indicator. The shrinkage strain could then be calculated by comparing current and initial readings. The specimens were also always installed in the same orientation to prevent any bias which might occur between the gauge studs and their corresponding sockets on the frame. The mass of the beams was recorded at each measurement interval as an indication of the rate of evaporation loss for each specimen.

### 3.7 Conclusion

Soil classified using ASTM D 2487-93 as a silty sand (SM) was tested to determine the optimal moisture content (OMC) for samples with a 5% cement content. An OMC of about 12% produced a peak dry density of 122.5 lb/ft<sup>3</sup>. The unconfined compressive strength (ASTM D 1633-00) of different mixture combinations was determined. Using strength criteria of 250+psi @ 7 days and 350-400psi @ 28-day, the optimal soil/cement/supplement ratios were selected. The splitting tensile strength (ASTM C 496-04) of select mixture combinations was also determined. The wetting and drying durability of the optimal mixtures for each type of supplemental material was determined following ASTM D 559-03. Finally, the shrinkage of each of the optimal mixtures was determined using a modified version of ASTM C 490-00. The results of these tests are analyzed in the next chapter.

## CHAPTER 4 ANALYSIS AND DISCUSSION

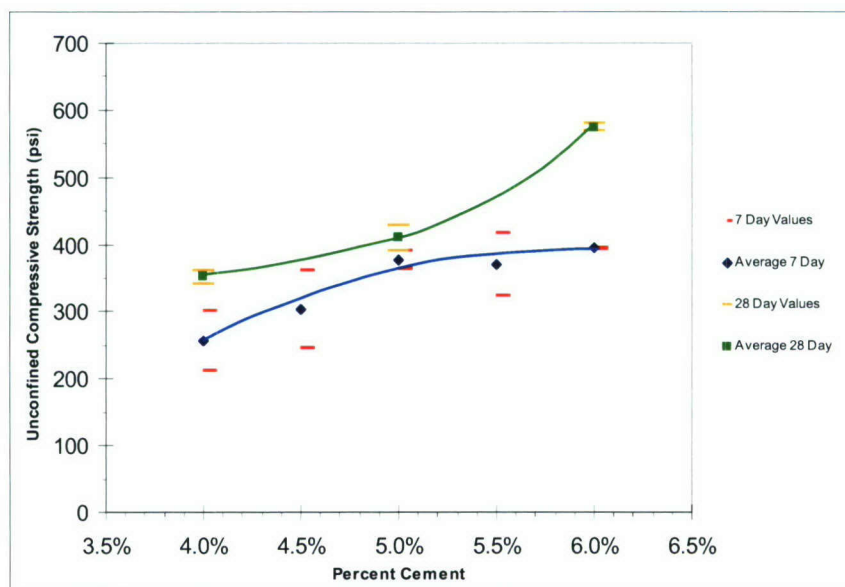
### 4.1 Introduction

Cylindrical samples of 4 inch diameter and 4.6 inch height having cement contents that range from 4-6% were molded. The first set of cylinders having cement contents from 4-6% by weight were tested for unconfined compressive strength. Specimens from the same initial set were also subjected to splitting tensile strength tests. The results of this testing served as the base values for evaluating mixtures altered with the addition of either Class F fly ash (FA), recycled carpet fiber (CF), or cement kiln duct (CKD). After selecting the optimal mix design for each type of supplemental material based on minimum 7 and 28-day strengths, two performance criteria were investigated. ASTM D559-03—*Standard Test Methods for Wetting and Drying Compacted Soil-Cement Mixtures* was performed to determine the durability of the different mixtures. A modified version of ASTM 490-00a—*Standard Practice for Use of Apparatus for the Determination of Length Change of Hardened Cement Paste, Mortar, and Concrete* was also performed to analyze the shrinkage experienced by each mixture combination.

### 4.2 Unconfined Compressive Strength

#### 4.2.1 Baseline Conditions— No Supplemental Materials

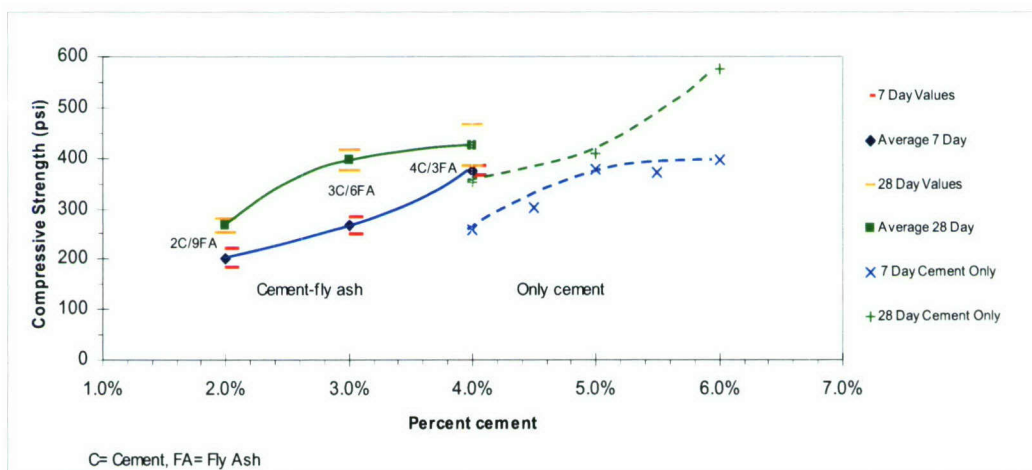
Optimal values were selected using the lowest minimum State DOT compressive strength requirements which was reported by the Louisiana DOT at 250 psi (George, 2001). Most other state DOTs have a higher minimum threshold in the range of 350-400 psi. The premise was that utilizing the lowest practical strength value would produce the most favorable distribution of shrinkage cracks (George 1971). Small closely spaced cracks are less likely to translate through asphalt overlays built on soil-cement bases. Higher tensile strength allows greater stresses to develop before cracking is initiated. This leads to larger more widely spaced cracks which can reflect through pavement. The results of method A of ASTM 1633-00— *Standard Test Methods for Compressive Strength of Molded Soil-Cement Cylinders* are shown in Figure 4-1. The 5% cement content sample's performance served as the bases for evaluating any potential benefits from incorporating any of the industrial byproducts. There is a higher rate of strength gain with increasing cement content noted for the 28-day cured samples. The data shows a wide variation in the 7-day strengths but only a small variation for 28-day strengths.



**Figure 4-1: Unconfined compressive strength vs. cement content.**

#### 4.2.2 Fly Ash

The 7 and 28-day strength values for the cement-fly ash mixtures are presented in Figure 4-2. The solid lines in Figure 4-2 represent the cement-fly ash combinations tested, while the dashed lines repeat the cement only data shown in Figure 4-1. The upper lines (28-day test results) each indicated that strength built as curing time increased. The pure cement strength values appear to be growing at a faster rate as cement content increased. The strength values were clearly enhanced by the presence of fly ash at the shared 4% cement level. The 7-day cured mixture with 4% cement and 3% fly ash appeared to be particularly strengthened when compared to the control mix at the same cement content.



**Figure 4-2: Unconfined compressive strength vs. percent cement. (FA admixture).**

#### 4.2.3 Carpet Fiber

The unconfined compressive strength (UCS) data for various soil-cement-carpet fiber combinations is presented in Figure 4-3. Cavey et al. (1995) reported difficulties with in-place mixing techniques used for incorporating recycled polypropylene fibers into soil-cement. With this type of equipment homogeneous mixtures could not be achieved, but dry batch mixing equipment designed for concrete proved successful. It is believed that

recycled carpet fibers are more difficult to incorporate into soil-cement than the recycled polypropylene fibers. It is therefore concluded that in-place mixing techniques are not appropriate for recycled carpet fiber. Murray et al. (2000) reported that after using a batch mixer, further hand mixing was required for carpet fiber to achieve a uniform distribution. The problem encountered using mechanical mixers with these fibers is the formation of low density clumps which float to the mixture's surface preventing the formation of a homogeneous mixture. This contrasts with a successful commercial application of recycled carpet fiber as a supplemental material for concrete documented by Wang et al. (2000). The fibers were added directly to the ready-mix truck before placement. The larger aggregate size and more viscous nature of concrete probably helped with fiber dispersion eliminating the need for hand mixing.

For this investigation most samples were compacted at a target moisture content of 12%. This was the optimal moisture content (OMC) of the 5% cement control sample. However, given that the inclusion of recycled carpet fibers was shown to reduce the effectiveness of the compaction process (Murray et al. 2000), a sample was compacted with a moisture level of 15%. The higher compressive strength of this sample was probably due to the extra lubrication provided by the moisture. These results were surprising because normally compaction at above OMC will reduce strength. The wicking action of the carpet fibers may have also created a hydrophilic environment where the extra moisture allowed for a more complete hydration of the cement.

The graph of UCS vs. cement content (Figure 4-3) has two solid lines which represent mixtures enhanced with carpet fibers. There is a linear trend for both the 28-day (upper line) and the 7-day values (lower line), but the individual specimen's strength values showed much more variation for the 28-day cured specimens. This contrasts with the data from the cement only specimens (Figure 4-1) which had greater variance in their 7-day data. The moisture variations caused by the wicking action of carpet fibers may have interfered with the curing process on lower strength samples.

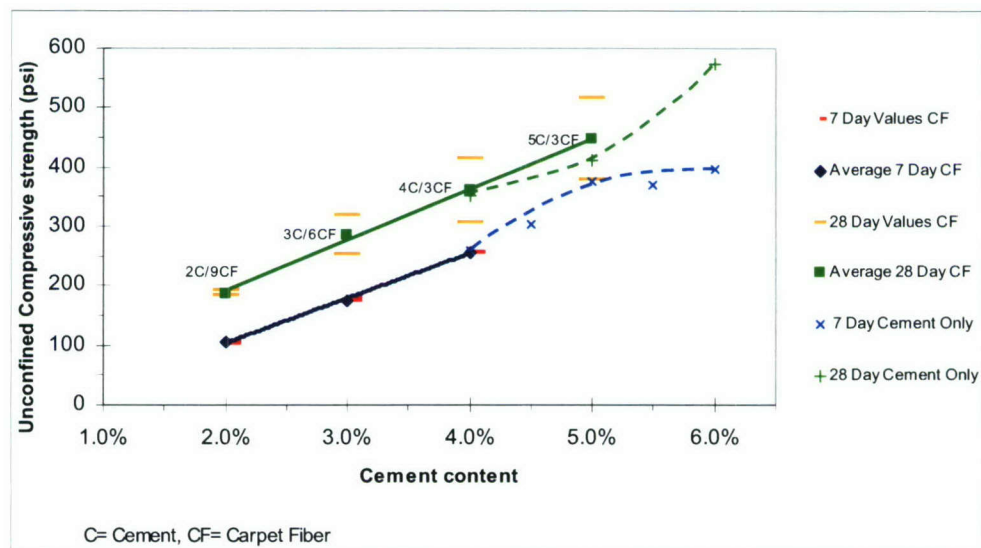
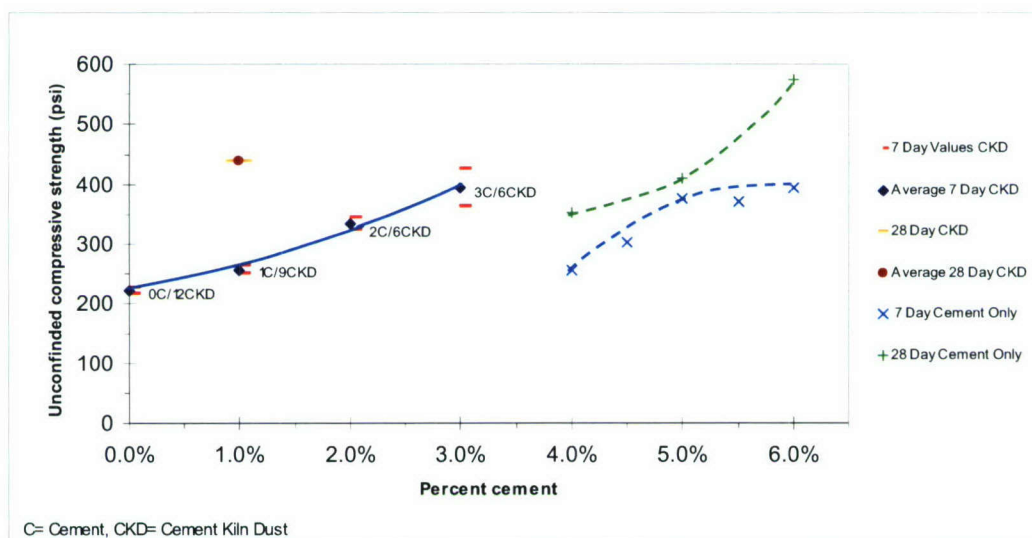


Figure 4-3: Unconfined compressive strength (UCS) vs. percent cement (cement-carpet fiber mixture).

#### 4.2.4 Cement Kiln Dust (CKD)

The unconfined compressive strength test results for soil-cement and CKD mixtures are presented in Figure 4-4. The addition of CKD allowed a significant reduction in the amount of cement required to meet the minimum strength criteria established for this test program. It appears necessary to include a small amount of cement (1%) to build-up the 7-day strength value. The two specimens molded with 12% CKD and 0% cement were not able to reach the required 250+ psi lower bound.

In Figure 4-4 the solid line represents the 7-day strength values for the low concentrations of cement combined with CKD. When compared to the lower dashed line (7-day values cement only) it is clear CKD is capable of rapidly building strength in excess of the levels cement alone can produce. The 7-day strength of 3C/6CKD is 395 psi which is almost as much as the 28-day strength of 5C/-- (409 psi). With the 3C/9CKD sample (data not graphed but included in Table 4-4) the 7-day strength (442 psi) exceeds the 28-day value for 5C/-- (409 psi).



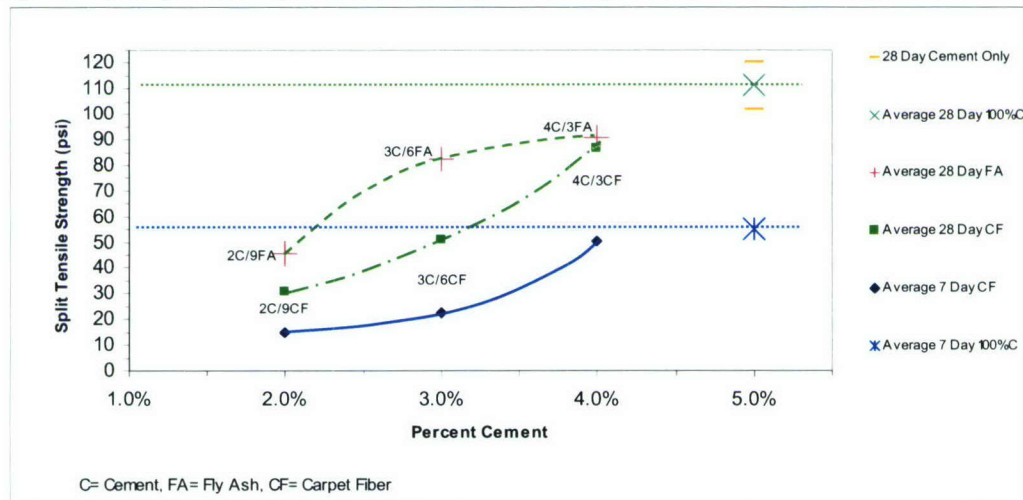
**Figure 4-4: Unconfined compressive strength vs. percent cement (cement-CKD mixture).**

#### 4.3 Splitting Tensile Strength

The indirect tensile strengths of selected mixtures were evaluated using ASTM C 496-04— *Standard Test Method for Splitting Tensile Strength of Cylindrical Concrete Specimens*. The results presented in Figure 4-5 indicate that none of the samples with supplemental materials were able to surpass the 5% cement base's strength. The 4C/3FA sample had a strength of 90 psi while the 4C/3CF sample yielded at 87 psi. Based on the compressive strength criteria, the optimal soil-cement-fly ash combination was 3C/6FA. The tensile strength of this mixture was only 82 psi, a 26% drop from the baseline value of 111 psi.

There were more variations in the strengths of individual specimens noted with higher concentrations of cement. In Figure 4-5 the upper dashed line represents the 28-day strength of soil-cement-fly ash combinations. The middle dash-dot-dash line represents the 28-day strength of soil-cement-carpet fiber. Included is a diagonal guideline which

projects a linear trend of increasing strength with increasing cement content. It appears that although the soil-cement-fly ash combinations are generally stronger at lower cement percentages, they are not gaining strength at the same rate as the mixtures with carpet fiber. At 7 days the 4C/3CF sample had nearly the same strength as the 5% baseline with no supplemental material (51 psi vs. 55 psi). This indicates an early enhancement by the presence of carpet fiber when the cement content was increased from 3% to 4%. The tensile strength performance of commercially available fibers was evaluated, and it was found that carpet fibers and new fibers provided about the same benefit. From the results of these two tests it was concluded that aside from the much more uniform mixing properties, recycled carpet fiber could provide equivalent benefits to commercial fibers.



**Figure 4-5: Splitting tensile strength vs. percent cement.**

#### 4.4 Wetting and Drying Durability

Specimens molded with the optimal supplemental material percentages were subjected to a durability test to evaluate their performance when exposed to repeated moisture extremes. The same test was conducted on similar specimens with cement only (control mix). The goal was to determine if these optimal mixtures can permanently hold a mass together with volume stability. ASTM D 559-03—*Standard Test Methods for Wetting and Drying Compacted Soil-Cement Mixtures* was performed to evaluate the durability of the different mixtures. The standard calls for 2 strokes with a force of 3lb to be applied to all exposed surfaces using a specified wire brush. As shown in Table 4-1 the maximum allowable mass loss is 14% for a soil type of SM (silty sand) according to the ACI (1990). Volume changes are limited to 2% of the initial volume, and the maximum moisture content cannot exceed the quantity required for saturation at the time of molding ACI (1990). The results of the testing are shown in Table 4-2.

Table 4-1: PCA allowable weight loss criteria (ACI 1990).

AASHTO soil group	Unified soil group	Maximum allowable weight loss, percent
A-1-a	GW, GP, GM, SW, SP, SM	14
A-1-b	GM, GP, SM, SP	14
A-2	GM, GC, SM, SC	14*
A-3	SP	14
A-4	CL, ML	10
A-5	ML, MH, CH	10
A-6	CL, CH	7
A-7	OH, MH, CH	7

\*10 percent is maximum allowable weight loss for A-2-6 and A-2-7 soils.

**Additional criteria**

1. Maximum volume changes during durability test should be less than 2 percent of the initial volume.
2. Maximum water content during the test should be less than the quantity required to saturate the sample at the time of molding.
3. Compressive strength should increase with age of specimen.
4. The cement content determined as adequate for pavement, using the PCA criteria above, will be adequate for soil-cement slope protection that is 5 ft or more below the minimum water elevation. For soil cement that is higher than that elevation, the cement content should be increased two percentage points.

All of the mixture combinations were able to pass the volume change and moisture content limitations. The low maximum volume change for the 3C/6FA specimen appears to be related to an unusually long specimen length which created a larger reference volume, combined with a low absolute loss in length. The higher maximum moisture content of the specimens with carpet fiber could be attributed to the higher porosity created by the inclusion of the fibers which often protrude from the specimens.

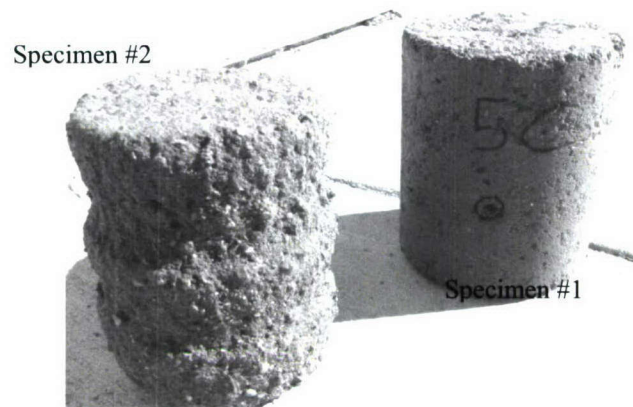
Table 4-2: Durability test results.

Mixture	Max $\square_d$ (lb/ft <sup>3</sup> )	$\Delta v_{\max}$	$\Delta \text{mass}_{\max}$ Spec. #1	$\Delta \text{mass}_{\max}$ Spec. #2	MC <sub>Max</sub>
5C / --	120.5	1.62%	2.6%	28.8%	9.22%
4C / 3FA	119.1	1.47%	2.7%	29.2%	9.60%
3C / 6FA	122.2	0.93%	3.4%	29.4%	9.10%
4 / 3CF	116.4	1.45%	3.5%	25.7%	11.02%
1C / 9CKD	120.0	1.65%	3.2%	26.3%	9.75%

(C= Cement, FA= Fly Ash, CF= Carpet Fiber, CKD= Cement Kiln Dust)

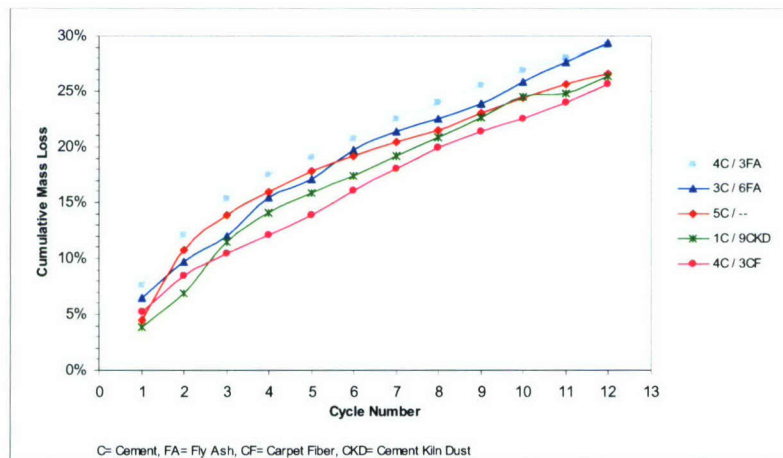
- $\Delta v_{\max}$ : maximum volume change-specimen 1 (unbrushed)  
 $\Delta \text{Mass}_{\max}$ : maximum mass loss-specimen 1 (unbrushed)  
 $\Delta \text{Mass}_{\max}$ : maximum mass loss-specimen 2 (brushed)  
 $\text{MC}_{\max}$ : maximum moisture content specimen 1 (unbrushed)

Specimen #2 in Figure 4-6 was subjected to 12 cycles of wetting and drying plus brushing with a standard wire brush. Its appearance obviously contrasts with specimen #1 which was only subjected to wetting and drying cycles.



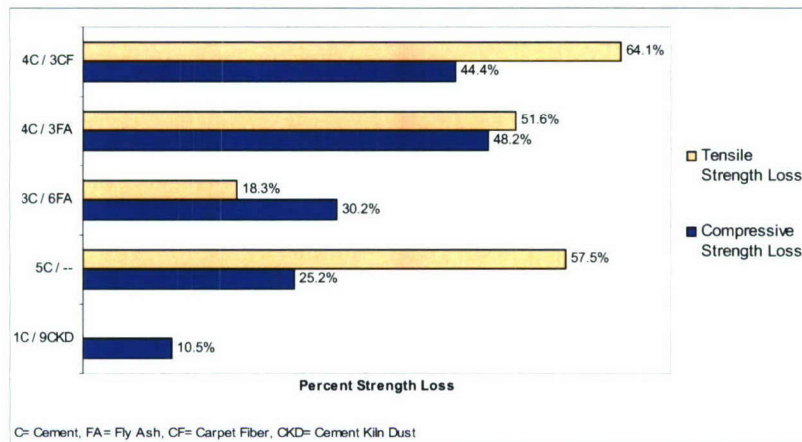
**Figure 4-61: Specimens used in durability test.**

The mass loss of specimen #2 versus brushing cycle number is shown in Figure 4-7. All of the samples exceeded the allowable 14% mass loss given by the PCA for silty sand soils (SM) (ACI 1990). After a higher initial rate of mass loss the different samples appear to loose a fairly consistent amount after each cycle. The best performer was the specimen with carpet fiber, but all of the cumulative losses were similar.



**Figure 4-7: Cumulative mass loss after each durability cycle (specimen #2).**

After finishing all 12 cycles of wetting/drying plus brushing, the minimum cross section of specimen #2 was determined and an unconfined compressive strength test was performed. This additional testing was an extension of the standard's procedure which gives another indication of durability. The splitting tensile strength was determined for specimen #1 which was only subjected to the moisture extremes with no brushing. The results are shown in Figure 4-8. CKD retained the largest portion of its original compressive strength and had the highest overall compressive and indirect tensile strengths after the durability cycles.

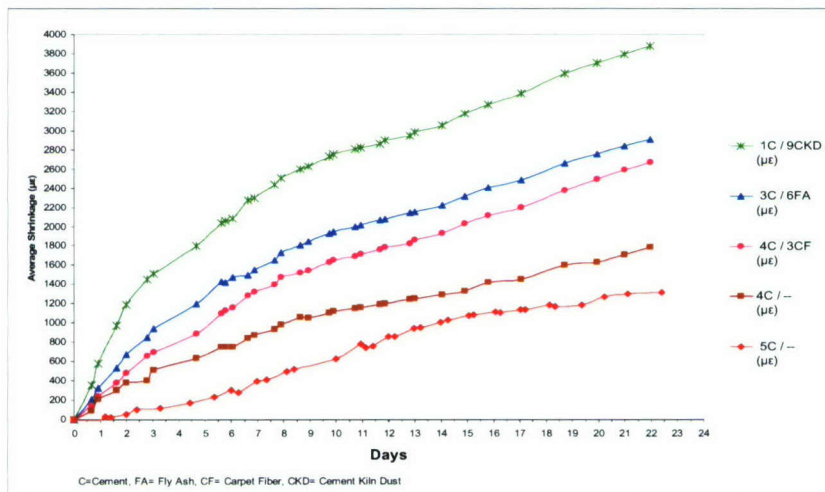


**Figure 4-82: Compressive and tensile strength losses after 12 wetting and drying cycles.**

#### 4.5 Shrinkage

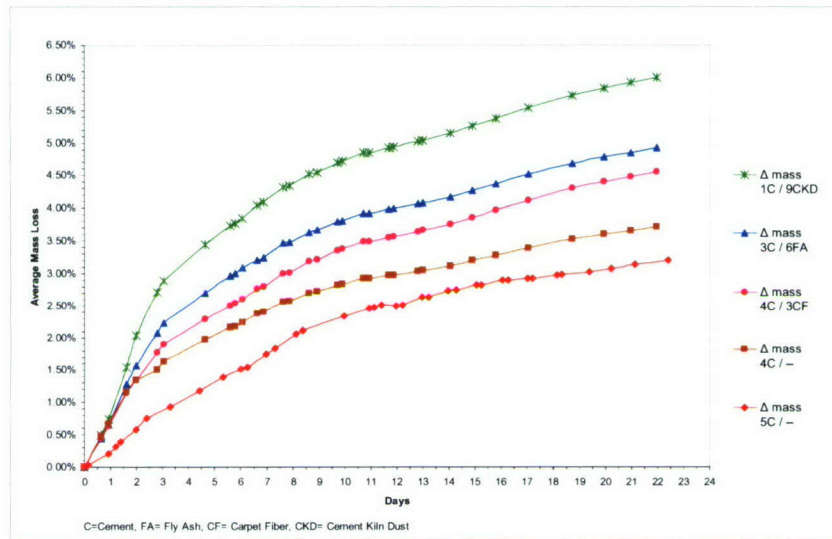
A modified version of the procedure outline in ASTM C 490-00 was performed to determine the length change of air dried samples. Corrections were made for the expansion and contraction of samples due to temperature changes and the mass was recorded with each length measurement.

It can be seen from both Figure 4-9 that the 1C/9CKD had both the highest initial strain rate and maximum strain. It also lost the most mass (Figure 4-10). It is believed this is largely due to the higher moisture used in molding the sample. Note that Norling (1973) concluded that changes in cement content, density, and temperature have much less affect on shrinkage than the initial compaction moisture content.



**Figure 4-9: Shrinkage as a function of time.**

The 5C/-- sample had the lowest initial rate and maximum strain as well as the smallest mass loss. This can be attributed to its low moisture content (MC) and high density. Evaporative losses are reduced when void spaces are eliminated at high densities because there are fewer pathways for moisture movement.



**Figure 4-10: Evaporative mass loss of shrinkage samples as a function of time.**

The 4C/-- sample was the second best performer with a low initial rate and maximum strain. Its small mass loss was also only surpassed by the 5/-- sample. The initial MC could not have been the only factor responsible for this performance. The 3C/6FA sample had a lower initial MC yet lost significantly more mass 4.9% compared to 3.7%. The fly ash sample also experienced a higher initial strain rate 274 (in/in)/day ( $\mu\epsilon$ ) versus 144 (in/in)/day ( $\mu\epsilon$ ). The maximum strains were also larger for the 3C/6FA sample. The dry densities of the 3C/6FA and 4C/-- samples were very close which probably negates density as a contributing factor in explaining the better performance of the cement sample. The 4C/3CF sample provides an interesting comparison with the 4C/-- sample. Adding the carpet fiber slightly raised the moisture content and slightly reduced the dry density but had a larger effect on the sample's shrinkage performance. Both the initial strain rate and maximum strain were raised by the carpet fiber's presence. This is the opposite of the desired effect and discourages the use of recycled carpet fiber as a supplement for soil-cement.

#### 4.6 Conclusion

Samples made with a 5% cement content served as the basis for evaluating any potential benefits from the industrial byproduct supplemental materials. The unconfined compressive strength of the 7-day 4C/3FA sample (373 psi) appeared to be particularly enhanced when compared to the control mix at the same cement content (257 psi). This may primarily be due to fly ash's ability to act as a filler of void spaces. It was found that recycled carpet fibers provided about the same compressive and splitting tensile performance as commercial fibers, but were more difficult to distribute uniformly. The

addition of CKD allowed a significant reduction in the amount of cement required to meet the minimum strength criteria established for this test program. It appears necessary to include a small amount of cement (1%) to build-up enough strength to exceed the target 7-day strength value (250+ psi).

None of the samples with supplemental materials were able to surpass the 5% cement base's splitting tensile strength at either 7 or 28 days. Although no particular tensile strength enhancement was observed as a result of the inclusion of carpet fibers, there was a noticeable improvement in ductility. All of the samples exceeded the allowable 14% mass loss given by the PCA (ACI 1990) for silty sand soils (SM). An aggressive brushing technique was the likely cause. The mass loss rates were comparable between the samples and the evaluation of the relative performance of each sample is believed to be valid. Specimens with CKD retained the largest portion of their original compressive strength and had the highest overall compressive and indirect tensile strengths after the durability cycles. The CKD specimen was also among the best performers with a small relative mass loss.

The 1C/9CKD sample had both the highest initial strain rate and maximum strain. It also lost the most mass through evaporation. It is believe this is largely due to the higher moisture content used in molding the sample. The 5% cement sample had the lowest initial rate and minimum shrinkage strain as well as the smallest mass loss. This may be attributable to its low moisture content (MC) and high density.

## **CHAPTER 5**

### **CONCLUSIONS AND RECOMMENDATIONS**

#### **5.1 Conclusions**

1. The compressive strengths of the samples were enhanced by the presence of fly ash. This was especially evident in the early stages of curing.
2. It was found that recycled carpet fibers provided about the same compressive and splitting tensile performance as commercial fibers. However they were difficult to distribute uniformly in soil-cement.
3. The addition of CKD allowed a significant reduction in the amount of cement required to meet the minimum compressive strength criteria established for this test program.
4. None of the samples with supplemental materials were able to surpass the 5% cement base's splitting tensile strength at either 7 or 28 days
5. Although there was no particular tensile strength enhancement observed as a result of the inclusion of carpet fibers, the post peak response was significantly altered. The addition of fibers produced a much more ductile failure.
6. All of the samples exceeded the allowable 14% mass loss given by the PCA for silty sand soils (SM) indicating probable durability problems.
7. Specimens with CKD retained the largest portion of their original compressive strength and had the highest overall compressive and indirect tensile strengths after the durability cycles. CKD specimen #2 was also among the best performers with a small relative mass loss.

8. Specimens with CKD had both the highest initial shrinkage rate and maximum shrinkage. They also lost the most mass through evaporation. It is believed this is primarily due to the higher moisture content used in molding the sample.
9. The 5% cement sample had the lowest initial rate and minimum shrinkage strain as well as the smallest mass loss. This may have been attributable to its low moisture content (MC) and high density.
10. Adding the carpet fibers increased the initial shrinkage rate and maximum shrinkage. This is the opposite of the desired effect and discourages the use of carpet fiber as a supplemental material for soil-cement to reduce shrinkage. This may be caused by the fibers increasing the sample's evaporation rate.

11.

## **5.2 Recommendations**

1. CKD could prove to be a useful supplement for soil-cement. The high shrinkage attributes are in large part due to the high moisture content at compaction. If evaporative losses could be controlled through the use of an emulsion sealer, for example, its shrinkage potential could be mitigated. Further study is warranted.
2. The addition of CKD to the research program was a late decision and inadequate supplies of soil and laboratory time prevented a more complete evaluation of its potential. Focused research is recommended to fully evaluate effects of compaction moisture content.
3. Durability testing on a similar soil type should be expanded to determine the lower bound of cement content which will pass the PCA criteria. Freeze-thaw testing (ASTM 560) could also provide additional insight on durability.
4. More research is needed with regard to the use of recycled carpet fiber in cement treated bases investigating its ability in bridging shrinkage cracking.
5. The effect of these waste recycle materials on cement-treated aggregate bases is recommended.

## **REFERENCES**

American Concrete Institute (ACI) (1990), *State-of-the-Art Report on Soil-Cement*, ACI 230.1R-90, Committee 230, Farmington Hills, MI.

Choi, Y.K., and Hansen, K.D. (2005), "RCC/Soil-Cement: What's the Difference?" *Journal of Materials in Civil Engineering*, Vol. 17, No.4, August, pp. 371-378

George, K.P. (2002), "Minimizing Cracking in Cement Treated Materials for Improved Performance," *Research and Development Bulletin No. RD123*, S Portland Cement Association Skokie, Ill

Gaspard, K.J. (2000), "Evaluation of Cement Treated Based Courses," *Technical Assistance Report Number 00-ITA*, Louisiana Transportation Research Center, Baton Rouge, LA

Arora, S. and Aydielk, A.H. (2005), "Class F Fly-Ash-Amended Soils as Highway Base Materials," *Journal of Materials in Civil Engineering*, Vol. 17, No. 6 November/December, pp. 640-649

Wang, Y., Wu, H.C., and Victor, C.L. (2000), "Concrete Reinforced with Recycled Fibers," *Journal of Materials in Civil Engineering* Vol. 12, No.4, November, pp. 314-319

Somayaji, S. (2001), *Civil Engineering Materials*, 2<sup>nd</sup> ed., Prentice-Hall, Inc., Upper Saddle River, NJ.

Sobhan, K. and Mashnad, M. (2002), "Tensile Strength and Toughness of Soil-Cement Fly-Ash Composite Reinforced with Recycled High Density Polyethylene Strips," *Journal of Materials in Civil Engineering*, Vol. 14, No.2, April pp. 177-184

Turner-Fairbanks Highway Research Center (TFHRC) (2002a), "Coal Fly Ash User Guideline—Stabilized Base," <http://www.tfhrc.gov/hnr20/recycle/waste/cfa55.htm>, Accessed on April 30, 2006

"Properties and Uses of Cement-Modified Soil", Portland Cement Association (PCA), 2003

George, K. P., "Soil Stabilization Field Trial", Interim Report, The University of Mississippi, 2001.

Norling, L.T., "Minimizing Reflective Cracks in Soil Cement Pavements: A Status Report of Laboratory Studies and Field Practices", Highway Research Record 442, Washington, D.C., 1973.

Little, D. N., T. Scullion, P.B.V.S. Kota and J. Bhuiyan, "Guidelines for Mixture Design of Stabilized Bases and Subgrades", FHWA/TX-45/ 1287-3 F, Texas Department of Transportation, Austin, Texas, October 1995.

Kota, P.B.V.S, T. Scullion and D.N. Little, "Investigation of Performance of Heavily Stabilized Bases in Houston, Texas, District", Transportation Research Record 1486, Washington, D.C., 1995.

George, K.P., "Crack Control in Cement-Treated Bases", Final Report, Civil Engineering Department, The University of Mississippi, 1970.

Colombier, G. and J.P. Marchand, "The Precracking of Pavement Underlays Incorporating Hydraulic Binders", Proceedings, Second International RILEM Conference on Reflective Cracking in Pavements, E and FN Spon, London, 1993.

Davidson, D.T., Katti, R. K., and Welch, D. E., "Use of Fly Ash with Portland Cement for Stabilization of Soil", Highway Research Board Bulletin 198, Washington, D.C., 1958.

Kasibati, K. and T. L. Conklin, "Field Performance Evaluation of Cement Treated

Bases with and without Fly Ash”, Transportation Research Record No. 1440, Washington, D.C., 1994.

“High-Volume Fly Ash Utilization Projects in the United States and Canada”, Publication No. CS-4446, Electric Power Research Institute Palo Alto, CA, February 1986.

Cavey, J. K., Krizek, R. J., Sobhan, K., and Baker, W. H. (1995), “Waste Fibers in Cement Stabilized Aggregate Base Course Material,” *Transportation Research Record*, Transportation Research Board, Washington, D. C. pp. 97-106.

Murray, J.J., Frost, J.D., and Wang, Y. (2000), “Behavior of a Sandy Silt Soil Reinforced with Discontinuous Recycled Fiber Inclusions,” *Transportation Research Record No. 1714*, Transportation Research Board, Washington, D. C. pp. 9-17

### **Publications and Presentations Related to Project**

#### **Abstracts**

*Rahim, Ashraf and Kirby, S.; ‘Industrial recycled waste materials as additives in soil-cement’; Presentation for the 1<sup>st</sup> International Conference of Environmental Management, Engineering, Planning and Economics; Greece, 2007.*

Soil-cement is a term used to describe a soil or aggregate that has been treated with a relatively small proportion of Portland cement. The objective of the treatment is to amend undesirable properties of problem soils or substandard materials so they are suitable for use in construction. Research was conducted to investigate the performance of a variety of lean soil-cement mixtures enhanced with industrial byproducts. Three industrial byproducts were evaluated in this research— Class F fly ash, recycled carpet fiber, and cement kiln dust (CKD). Historically, Class C fly ash has been the preferred supplemental material for concrete because of its self-cementitious properties, but Class F fly ash is produced in larger quantities and finds fewer commercial applications. The source of fly ash is coal-fired electric power plants where fly ash is produced as a combustion byproduct formed as flue gasses cool. According to the American Coal Ash Association, as of 2001, 58% of the coal combustion products were Class F fly ash and 68% of these required disposal for lack of use. Approximately 3% of the municipal solid waste stream generated in the United States is composed of carpets and textiles. Carpet byproducts including production waste and recycled post-consumer carpet are estimated to be produced at a rate of 2,000,000 tons/year. Cement Kiln Dust (CKD) is a byproduct of the cement production process. CKD is captured by air pollution control equipment installed on the chimneys of cement kilns. According to the EPA in 1990, the cement industry generated an estimated 12.7 million metric tons of CKD at 111 plants in 38 states of which 4 million metric tons were disposed of in piles, quarries, and landfills. Criteria based on the 7 and 28-day compressive strengths were adopted, on which each optimal supplemental material percentage was selected (250 psi @ 7 days and 350-400 psi @ 28 days). A series of unconfined compressive strength tests were performed along with splitting tensile strength tests. ASTM D 559-03 (wetting and drying durability) was

conducted on supplemental material combinations which met the strength criteria with the minimal cement content. The addition of Fly Ash enhanced the compressive strength. Addition of CKD allowed a significant reduction in the amount of cement required to meet the minimum compressive strength criteria established for this test program. The addition of carpet fibers produced a much more ductile failure in the splitting test. Specimens with CKD retained the largest portion of both compressive and tensile after the durability cycles.

**Modeling the strong electro-optical response of “de Vries”  
ferroelectric liquid crystals**

Project Investigator:

Karl Saunders  
Physics Department  
California Polytechnic State University  
San Luis Obispo, California

### Overall project aim

The eventual goal is a model for the strong electro-optical response that has been recently observed in a novel class of ferroelectric liquid crystals, known as de Vries liquid crystals. While there have been substantial experimental studies of these new de Vries liquid crystals, a rigorous theoretical study has not, until now, been undertaken. The first stage of this project is the development and analysis of a model for non-chiral de Vries liquid crystals. This model can then form the basis for a more general model that incorporates both chiral and non-chiral liquid crystals.

The project has to date been very successful. A manuscript detailing the design, analysis and results of our non-chiral model is currently being prepared for submission to *Physical Review Letters*, a top-tier physics journal. Two of the coauthors are Cal Poly undergraduate students who worked on the project throughout the summer.

### 1) Summary

#### a) Results to Date

We have developed and analyzed a comprehensive new model for non-chiral liquid crystals near the AC transition. We have also found a regime in which the model predicts behavior consistent with the development of a non-chiral de Vries structure. This result is scientifically and technologically significant as it supports the theory that the de Vries smectic is an unusual type of smectic-A phase rather than a new distinct phase. It also means that de Vries liquid crystals can be described and studied using traditional liquid crystal theory.

Additionally, this is the first rigorous analytic theory to predict *and* explain this unusual type of behavior, and it allows several universal features of non-chiral de Vries liquid crystals to be identified. These features are consistent with experimental studies that have been carried out by other scientists. From a technological perspective this makes our model exciting as it means that it can (and will) be used as a foundation for the development of a more general model that incorporates both chiral and non-chiral liquid crystals. This model will then be used to make concrete predictions and evaluations of the electro-optical response of the materials. This is crucial in assessing the feasibility of these materials for use in new liquid crystal display technology.

We had originally proposed that a parallel theory of the critical fluctuations at the AC transition could go some way to explaining the development of the de Vries structure. However, our model based on Landau theory proved to be unexpectedly rich, complex and adaptable. It by itself provides a very satisfactory explanation that correlates extremely well with the experimental results. The theory using critical fluctuations is still of interest but will be undertaken at a later date, given our priority of generalizing our successful Landau model to include chirality.

In addition to the de Vries structure, our non-chiral model predicts a range of liquid crystal behavior, including some new types of phase transitions. These unexpected, yet very welcome, features make our non-chiral model very promising for the study of other

liquid crystal phases and phase transitions. There are several particular aspects that we feel merit further exploration at a later date.

In a particularly exciting development, a highly regarded liquid crystal experimentalist, Professor Jonathan Fernsler has recently joined the department. In addition to being an accomplished liquid crystal experimentalist, Professor Fernsler has a particular interest and expertise in de Vries materials. We have already been discussing the modeling process with Professor Fernsler and are planning a close and active collaboration.

b) Summary of Plans for continuation of the project

It was originally proposed that the development and study of the chiral model would be completed by the project end date. However, as discussed above, unexpected richness and complexity of the non-chiral model merited a more in-depth and rigorous analysis. Such is the importance of our results of this stage, that we are submitting an article on the non-chiral model alone.

We now propose to generalize this model to incorporate chiral liquid crystals. Incorporating chirality into the model is a non-trivial process but we are confident that we can develop a suitable chiral model given the success of the underlying non-chiral model. Having designed the more general chiral model we will then apply the same analytical procedures as we did for the non-chiral model. Once we have established the existence of a regime in our model with behavior consistent with the development of a de Vries structure, we will analyze the electro-optical properties. Our plans for this project are presented in more detail in a new proposal “Modeling Chiral de Vries Liquid Crystals and their Electro-Optical properties”.

As discussed above, we will collaborate closely with Professor Fernsler on this project. Professor Fernsler comes to Cal Poly with a wealth of experience, expertise and data on chiral de Vries liquid crystals. We have begun planning with Professor Fernsler for a theoretical-experimental collaborative study. The combination of theoretical and experimental expertise in such a collaboration would have incredible scientific value. It would also form the basis for the growth of a Cal Poly Liquid Crystal group.

2) Technical details and specifics of results:

a) Results pertaining to de Vries materials

The Landau theory we have developed for non-chiral liquid crystals keeps track of the development and interplay of three types of liquid crystal order: orientational, layering, and tilting. This is the first model to keep track of all three types of order. Analysis of our model predicts that, in certain regimes, the smectic A phase will exhibit behavior consistent with the development of a de Vries structure. The three main results supporting this are as follows:

- Orientalional Order: We have shown that within a certain regime, the orientational order of the smectic A phase will begin to *decrease* as the smectic C phase is approached and will then begin to increase again as the C phase is

entered. Our model is the first to predict such highly unusual behavior. Indeed, this is the only situation we know of in which a material becomes less ordered as it approaches a more ordered phase, and then becomes more ordered once the phase is entered. As such seems to be a unique signature of the de Vries phase. The behavior for the orientational order predicted by our model corresponds very well with existing experimental results.

- Layer Spacing: We have also shown that in the same regime, the layer spacing of the A phase should increase as the de Vries structure develops. This also agrees with experimental observation. That this somewhat counter intuitive feature of the de Vries structure is predicted by our model makes our theory all the more compelling. It has been observed experimentally that the layer spacing changes very little once the C phase is entered. This feature makes de Vries liquid crystal materials appealing for use in liquid crystal displays, as it eliminates the problem of layer buckling and striping in the display. Our model predicts that in the regime of interest, the layer spacing does indeed remain virtually unchanged as the transition to the C phase is made.
- Importance of Isotropic-Smectic A transition: In general a variety of phase sequences leading up to the smectic C phase are possible for different liquid crystals. Our model also indicates that the development of the de Vries structure is favored by systems that exhibit a particular phase sequence: isotropic-smectic A-smectic C. This is also a general feature of all experimentally observed de Vries materials.

The development of a model that predicts these unusual and experimentally consistent features is in itself a scientifically significant result. The third feature also gives valuable insight into the mechanism for the formation of the de Vries structure. We believe that in de Vries materials it is the layering order that induces the orientational order of the smectic A phase. This is evidenced by the absence of a nematic phase (which possesses only orientational order) in such materials. Using our model we have shown that the transition from the A phase to the C phase occurs via a decrease in the coupling between the layering and orientational order. In a material in which the layering order is responsible for inducing the orientational order, the decrease in the coupling between the two orders will result in a decrease of orientational order- the unique feature of de Vries materials.

#### b) Additional results

In addition to predicting both standard liquid crystal behavior and de Vries behavior (in different parameter regimes), our model has other interesting and novel features:

- The existence of a first order smectic A to smectic A transition (in which the amount of order in the A phase changes discontinuously) and an associated critical end point.
- The existence of a first order nematic to smectic A phase boundary and the existence of a tri-critical point separating this first order boundary from a second order boundary.
- The possibility of a spontaneous development of chirality.

These features will be investigated and evaluated for possible future study.

### 3) Administrative Details

- A Dell 810 was purchased for \$2865 in December 2005 and has been used for the analysis of the non-chiral model.
- The PI took 4 units release time in Spring quarter.
- Two students were hired beginning Spring quarter. They worked on the analysis of the model and will appear as co-authors on the manuscript being submitted to Physical Review Letters.

### Publications and Presentations Related to Project

- Full publications  
Copies of full publications follow.

**Development of a micro system actuator employing shape memory  
alloys for exploring the nanotechnology world**

Project Investigator:

Richard N. Savage  
Materials Engineering  
California Polytechnic State University  
San Luis Obispo, California

# Disordering to Order: de Vries behavior from a Landau theory for smectics

Karl Saunders <sup>1\*</sup>, Daniel Hernandez <sup>1</sup>, and Staci Pearson <sup>1</sup>, John Toner<sup>2</sup>

<sup>1</sup>*Department of Physics, California Polytechnic State University, San Luis Obispo, CA 93407, USA and*

<sup>2</sup>*Department of Physics and Institute of Theoretical Science, University of Oregon, Eugene, OR 97403, USA*

(Dated: April 7, 2007)

We show that Landau theory for the isotropic ( $I$ ), nematic ( $N$ ), smectic  $A$ , and smectic  $C$  phases generically, but not ubiquitously, implies “de Vries” behavior: i.e., a continuous  $AC$  transition can occur with little layer contraction while the birefringence increases significantly once the system moves into the  $C$  phase. Our theory shows that 1st order  $AC$  transitions are also possible. These transitions can be de Vries-like, but in general need not be. Generally, de Vries behavior occurs in models with unusually small orientational order, and is preceded by a first order  $I - A$  transition. These results correspond well with experimental work to date.

Recently, an unusual new class of liquid crystals known as “de Vries smectics” [1] has drawn interest. They possess two defining features. Firstly, there is little change with temperature  $T$  of the layer spacing  $d(T)$  upon entry to the  $C$  phase, in contrast to the rapid geometrical contraction  $d(T) \propto \cos \theta(T)$  expected if the molecules tilt by a strongly temperature angle  $\theta(T)$ . Secondly, the birefringence increases significantly upon entering the  $C$  phase from the  $A$  phase. In fact, for de Vries materials with a 2nd order  $AC$  transition, the birefringence is seen[2, 3] to *decrease* with decreasing temperature as the  $AC$  transition is approached, reaching a minimum at the  $AC$  transition. This is the first example known to us of *decreasing* order as a lower symmetry phase is approached. For de Vries materials with a 1st order transition the birefringence increases moderately as the  $AC$  transition is approached and then jumps significantly at the transition[4]. Generally, de Vries smectics exhibit the phase sequence  $I-A-C$ , without a nematic phase. First order  $AC$  transitions are not always de Vries-like [4], contrary to some recent speculations.

In de Vries’ “diffuse cone model” [5] of these materials, the molecules “pre-tilt” in the  $A$  phase as the  $AC$  transition is approached, but in azimuthally random directions (hence reducing orientational order), so that there is no long range order in the tilting. Upon entering the  $C$  phase the molecules azimuthally order (hence increasing orientational order) without the significant layer contraction that occurs in conventional smectics whose molecules tilt at the  $AC$  transition.

In this Letter we show that in a complete, nonchiral Landau mean field theory for the isotropic ( $I$ ), nematic ( $N$ ),  $A$  and  $C$  phases, in which all three order parameters (orientational, layering, azimuthal tilt) and the layer spacing are coupled, de Vries behavior occurs in a finite fraction of parameter space, while other regions exhibit conventional behavior. The mean field phase diagram for our model is shown in Fig. 1. Here  $t_s$  and  $t_n$  are Landau theory parameters that control layering, and orientational order, respectively. We find that two main features

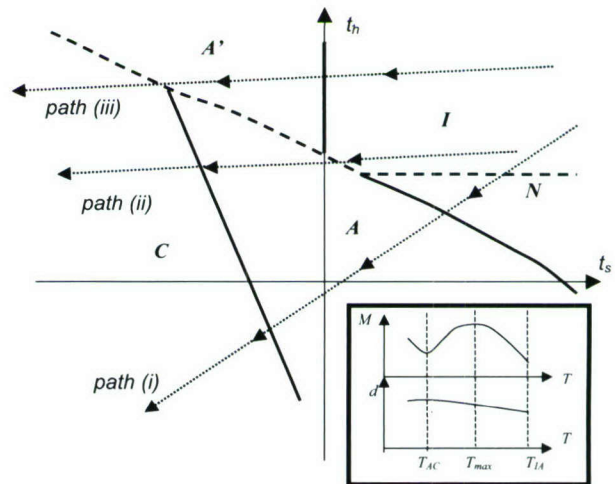


FIG. 1: The phase diagram in  $t_s$ - $t_n$  space for the  $I$ ,  $N$ ,  $A$ ,  $C$  phases. 1st and 2nd order phase boundaries are shown as dashed and solid lines respectively. Three decreasing temperature paths from the  $I$  to  $C$  phase are shown. Path (i) corresponds to a conventional material that does not exhibit de Vries behavior. Path (ii) corresponds to a material exhibiting de Vries behavior and a 2nd order  $AC$  transition. Path (iii) leads to a 1st order  $AC$  transition with de Vries behavior. The inset shows a possible temperature dependence of birefringence ( $M$ ) and layer spacing ( $d$ ) for path (ii).

are necessary for de Vries behavior- an unusually weak coupling between layering and orientational order, and a virtually temperature independent  $t_n$ . This latter feature would correspond to almost perfect excluded volume short range repulsive molecular interactions. Our theory also predicts that materials exhibiting de Vries behavior will almost always follow the phase sequence  $I-A-C$ . For 2nd order  $AC$  transitions we show that systems with an athermal  $t_n$  will exhibit the unusual feature of a decreasing birefringence as the transition is approached from the  $A$  phase. In fact, we predict that in the  $A$  phase, some systems will exhibit a birefringence that has a nonmonotonic dependence on temperature, as shown in the inset

of Fig. 1 [6]. Lastly, we also show that 1st order  $AC$  transitions can (but generally need not) be de Vries-like.

Another result of our theory, unrelated to de Vries behavior, is that this phase diagram predicts *two*, distinct smectic  $A$  phases of identical symmetry, denoted  $A$  and  $A'$ , separated by a first order phase transition. While 1st order transitions between two smectic  $A$  phases due to competition between two different layer spacings[7] have been predicted, our result shows that even without such competition,  $A$ - $A'$  transitions occur quite naturally, and should be far more common than was previously thought.

Any experiment in which temperature is varied at fixed concentration traces a path through this phase diagram. As usual in Landau theory, we assume throughout this paper that  $t_s$  and  $t_n$  are monotonically increasing functions of temperature; hence, as temperature is lowered, one moves monotonically from upper right to lower left in Fig. 1. Three qualitatively distinct paths of this type are shown. Path (i) is a typical path for a material that does *not* display de Vries behavior; along it,  $t_s$  and  $t_n$  both depend strongly on temperature. Paths (ii) and (iii) correspond to de Vries behavior with 2nd and 1st order  $AC$  and  $A'C$  transitions, respectively. Both paths have strongly varying  $t_s$  and weakly varying  $t_n$ ; i.e,  $t_n$  is virtually athermal. This would be the case if the  $IN$  transition is driven by a steric mechanism for which  $t_n$  depends strongly on concentration and weakly on  $T$ . We find that de Vries behavior occurs in this case, for sufficiently weak coupling between layering and orientational orders. It should also be pointed out that a path above (iii), away from the  $AA'C$  critical end point (CEP), would exhibit a 1st order  $A'C$  transition *without* de Vries behavior.

This restriction to nearly horizontal paths implies that de Vries systems should very rarely exhibit an  $N$  phase between the  $A$  and  $C$  phases, since, to cross the  $IN$  boundary, a nearly horizontal path in Fig. 1 would have to be “fine tuned” to start very close to the boundary. The most likely paths to see de Vries behavior are those like (ii) or (iii), showing phase sequence  $I$ - $A$ - $C$  or  $I$ - $A'$ - $C$ , in good agreement with experimental work to date.

In this Letter we focus on path (ii) and briefly discuss path (iii) at the end. The inset in Fig. 1 shows predictions for the layer spacing  $d(T)$  and birefringence  $M(T)$  as  $T$  is varied along path (ii). The *increase* of the layer spacing in the  $A$  phase as the  $AC$  transition is approached, though contrary to the de Vries picture of “pre-tilting” in the  $A$  phase, is seen experimentally [2]. In the  $A$ -phase, our Landau theory predicts:

$$M_A(T) = M_{max} - M_2 \left[ \frac{t_s(T) - t_s(T_{max})}{t_s(T_{max})} \right]^2, \quad (1)$$

where  $M_{max}$ ,  $M_2$  and  $T_{max}$  are positive constants. We have shown that if  $t_n$  is completely athermal,  $T_{AC} < T_{max}$ , where  $T_{AC}$  is the  $AC$  transition temperature, so that in systems that follow a path like (ii),  $M_A(T)$  will

decrease as the  $AC$  transition is approached. A subset of such systems will have  $T_{max} < T_{IA}$ , where  $T_{IA}$  is the  $IA$  transition temperature, so that  $M_A$  has a maximum *within* the  $A$  phase, as shown in the inset in Fig. 1. If  $t_s$  is linear in  $T$  in the  $A$ -phase (as expected for small  $T_{IA} - T_{AC}$ ), then  $M_A(T)$  will be *perfectly* parabolic in  $T$ .

In the  $C$  phase near the  $AC$  transition the critical temperature dependences of  $M$  and the tilt angle  $\theta(T)$  predicted by our Landau theory are:  $M$  linear in  $T$ , and  $\theta \propto (T_{AC} - T)^{\frac{1}{2}}$ . When fluctuation effects are included, we expect[8]  $\theta \propto (T_{AC} - T)^\beta$ , where  $\beta \approx 0.35$  is the order parameter critical exponent for the 3D XY model. The layer spacing scales with temperature according to:

$$d = d_0 + a(T_{AC} - T) + b(T_{AC} - T)^3, \quad (2)$$

where  $d_0$  is the value of the layer spacing at the  $AC$  transition and  $a$  and  $b$  are constants that depend on the Landau theory parameters. Clearly, if  $a$  is sufficiently small, which we find is the case in the  $C$  phase for sufficiently weak coupling between layering and orientational order, the layer spacing shows very little variation with temperature near  $T_{AC}$ . Significantly, we find that in cases like path (ii), the criterion for de Vries behavior of  $d(T)$  differs from that for  $M(T)$ . Hence, we predict that some systems will exhibit de Vries behavior of the birefringence, but *not* de Vries behavior of the layer spacing.

How do fluctuations affect our mean field Landau theory? While our theory predicts that the  $I$ - $A'$  transition is continuous, it is known[9] that fluctuations *always* drive the  $I$ - $A'$  transition first order, albeit only weakly so if fluctuations are small. Fluctuation effects will also shift the positions of *all* of the transitions we've found. We expect, however, that the topology and essential geometry of the phase diagram Fig. 1 should occur in real systems. The only qualitative difference we expect is that the  $I$ - $A$ - $A'$  critical end point (CEP) predicted by Landau theory will be replaced by an  $I$ - $A$ - $A'$  triple point.

In other regions of parameter space, our model has an  $NAC$  point; we will discuss this elsewhere[10].

In summary, de Vries behavior emerges quite naturally from our Landau theory. Equally importantly, conventional behavior (for both 1st and 2nd order transitions) also generically occurs for different Landau parameters. Thus, the model can accommodate all observed behaviors in *all* systems, and also predicts new behaviors not yet seen experimentally, like the first order  $A$ - $A'$  transition.

We will now briefly describe the formulation and analysis of our theory. A Landau theory for all four phases ( $I$ ,  $N$ ,  $A$ ,  $C$ ) must include order parameters for three types of order: uniaxial orientational order, tilt (azimuthal) order, and layering order. The first two are embodied in the usual second rank tensor orientational order parameter  $Q$ . The complex layering order parameter  $\psi$  is defined via the density  $\rho = \rho_0 + \Re(\psi e^{i\mathbf{q}\cdot\mathbf{r}})$  with  $\rho_0$  constant and  $\mathbf{q}$  the layering wavevector. Taking both  $\psi$  and  $Q$  to be spatially uniform implies the Landau free energy density

$f = f_Q + f_\psi + f_c$ , with the orientational ( $f_Q$ ) [11], layering ( $f_\psi$ ), and coupling ( $f_c$ ) terms given by:

$$f_Q = \frac{t_n \text{Tr}(\mathcal{Q}^2)}{12} - \frac{w \text{Tr}(\mathcal{Q}^3)}{18} + \frac{u_n (\text{Tr}(\mathcal{Q}^2))^2}{144}, \quad (3)$$

$$f_\psi = \frac{1}{2} t_s |\Psi|^2 + \frac{1}{4} u_s |\Psi|^4 + \frac{1}{2} K (q^2 - q_0^2)^2 |\Psi|^2, \quad (4)$$

$$f_c = \frac{1}{2} \left[ -g_1(q^2) q_i q_j Q_{ij} + g_2 q_i q_j Q_{ik} Q_{jk} \right] |\Psi|^2 + \frac{h}{4} (q_i q_j Q_{ij})^2 |\Psi|^4, \quad (5)$$

where  $q_i$  is the component of the layering wavevector in the  $i$ th direction, and the Einstein summation convention is implied. The constant  $q_0$  is the magnitude that the layering wavevector would have in the absence of coupling between layering and orientational order. For weak coupling,  $q \approx q_0$  and we Taylor expand  $g_1(q^2) \approx g_{10} + g'_1(q^2 - q_0^2)$ , where  $g_{10} \equiv g_1(q_0^2)$ , and  $g'_1 \equiv \frac{dg_1}{dq^2} \big|_{q^2=q_0^2}$ . The relatively large number of parameters in  $f$  is inevitable given the fact that the theory incorporates three types of order. Furthermore, this theory allows for the possibility of conventional or de Vries behavior. However, we will show that in general, any of the five possible phases shown in Fig. 1 can be accessed by allowing, at most, two of the above parameters ( $t_n$  and  $t_s$ ) to vary (with temperature). The presence or absence of the two defining features of de Vries behavior, small change in layer spacing and decreasing birefringence, depend only on the size of  $g'_1$  and the athermal nature of  $t_n$ , respectively. The remaining fixed positive parameters are required to stabilize the phases [12–14].

Choosing the arbitrary direction of the layer normals to be  $z$ , we seek the configuration of  $\psi$ ,  $\mathcal{Q}$  and  $\mathbf{q} = q\hat{\mathbf{z}}$  that minimizes  $f$ . The form of  $\mathcal{Q}$  that does so [10] is given by:

$$Q_{ij} = (-S + \sqrt{3}\eta) e_{1i} e_{1j} + (-S - \sqrt{3}\eta) e_{2i} e_{2j} + (2S) e_{3i} e_{3j}, \quad (6)$$

where  $\hat{\mathbf{e}}_3 = \mathbf{c} + \sqrt{1 - c^2} \hat{\mathbf{z}}$  is the average direction of the molecules' long axes, (i.e., the director). Here, in either smectic phase,  $\hat{\mathbf{z}}$  is normal to the layers; in the  $N$  and  $I$  phases the direction of  $\hat{\mathbf{z}}$  is arbitrary. The projection,  $\mathbf{c}$ , of the director onto the layers is the order parameter for the  $C$  phase. The other two principal axes of  $\mathcal{Q}$  are given by  $\hat{\mathbf{e}}_1 = \hat{\mathbf{z}} \times \hat{\mathbf{c}}$  and  $\hat{\mathbf{e}}_2 = \sqrt{1 - c^2} \hat{\mathbf{c}} - c\hat{\mathbf{z}}$ .  $S$  and  $\eta$  are proportional to the birefringence and biaxiality of the system, respectively. The  $A$  phase is untilted ( $\mathbf{c} = \mathbf{0}$ ) and uniaxial ( $\eta = 0$ ), while the  $C$  phase is tilted ( $\mathbf{c} \neq \mathbf{0}$ ) and biaxial ( $\eta \neq 0$ ). It is convenient to make the change of variables  $S = M \cos(\alpha)$  and  $\eta = M \sin(\alpha)$ . In the  $A$  phase,  $M$  is proportional to the birefringence.

We next minimize the free energy  $F$  over the variables  $M, \alpha, c, |\Psi|$  and  $q$ . Four qualitatively different types of minima are possible, corresponding to the four different symmetry phases ( $I, N, A, C$ ). Specifically, the  $I$  phase has  $M = 0; \Psi = 0$ ; the  $N$  phase has  $\Psi = 0, \mathbf{c} = \mathbf{0}$ , and  $\alpha = 0$ , but  $M \neq 0$ ; the  $A$  phase has  $\Psi \neq 0$  and  $M \neq 0$ , but  $\mathbf{c} = \mathbf{0}$ , and  $\alpha = 0$ ; and the  $C$  phase has all of the variables  $M, \alpha, \mathbf{c}$ , and  $\Psi \neq 0$ . We render minimization analytically tractable by assuming that the coupling term Eq. (5) is small, and treating it perturbatively. Standard phase transition analysis [15] leads to the phase diagram shown in Fig. 1. Equations for the loci of the phase boundaries are given in [16]. The minimization of our Landau free energy also leads to predictions for the temperature dependences of  $M$  and  $q$ . We find, in the  $A$ -phase,

$$M_A = M_0(t_n) + \frac{q_0^2 \Psi_0^2}{\gamma} (-3g_2 M_0(t_n) + \delta), \quad (7)$$

$$q_A^2 = q_0^2 + \frac{M_0(t_n)}{K} (-g_2 M_0(t_n) + g'_1 q_0^2 + \delta), \quad (8)$$

where  $\gamma \equiv w M_0(t_n) - 2t_n > 0$  and  $M_0(t_n) = (w + \sqrt{w^2 - 4u_n t_n})/2u_n$  is the “bare” value of  $M$ , i.e. its value in the absence of coupling. Likewise,  $\Psi_0 = \sqrt{-t_s/u_s}$  is the bare value of  $\Psi$ . For strongly  $T$  dependent  $t_s$  and athermal  $t_n$  the quantity  $\delta$  is most usefully expressed as

$$\delta(t_s, t_n) \equiv \alpha(t_s - t_s^{AC}), \quad (9)$$

where  $\alpha = (2hq_0^2 M_0)/u_s$  and  $t_s^{AC} = -(g_{10} - g_2 M_0)/\alpha$  is the value of  $t_s$  where  $\delta$  vanishes and the 2nd order  $AC$  transition occurs. In the  $A$  phase,  $\delta > 0$  and in the  $C$  phase  $\delta < 0$ . In the  $C$  phase we find

$$M_C = M_0(t_n) + \frac{q_0^2 \Psi_0^2}{\gamma} \left( -3g_2 M_0(t_n) - \frac{g_2}{2hq_0^2 \Psi_0^2} \delta \right), \quad (10)$$

$$q_C^2 = q_A^2(t_s^{AC}) + \frac{g'_1}{2Kh\Psi_0^2} \delta, \quad (11)$$

where a decrease in layer spacing requires  $g'_1 < 0$ . Finally, in the  $A'$ -phase, we find:

$$M'_A = \frac{q_0^2 \Psi_0^2 g_{10}}{t_n}, \quad (12)$$

$$q_A'^2 = q_0^2 \left( 1 + \frac{2g_{10}^2 \Psi_0^2}{K t_n} \right). \quad (13)$$

These results imply de Vries behavior for both birefringence and layer spacing. For a nearly horizontal experimental locus like path (ii) through the  $A$  phase, the  $T$  dependence of  $M$  (and hence birefringence) in Eq. (7) comes from the linear  $t_s$ -dependence of each of  $\Psi_0^2$  and  $\delta$  in the correction due to the coupling of layering and orientational orders. From Eq. (7) we see that non-monotonicity of  $M_A$  is due to a competition between the layering order,  $\Psi_0^2$  and the coupling  $\delta$  which increase and

decrease respectively as the  $AC$  transition is approached. This happens because, as the system moves deeper into the  $A$  phase, the layering order increases, thereby augmenting the weak orientational order due to the coupling between the two. However, as the  $AC$  transition (where at the director tilts away from the layer normal) is approached, this coupling necessarily decreases, and, hence, so does  $M_A$  which means the birefringence *decreases* as the  $AC$  transition is approached. Once in the  $C$  phase the birefringence begins to *increase* with decreasing  $t_s$ . It is straightforward to show that if  $g_{10} > 4g_2M_0$  then  $M_A$  exhibits a maximum inside the  $A$  phase, as shown in the inset in Fig. 1.

As the  $AC$  transition is approached within the  $A$  phase,  $q_A$  monotonically decreases and hence  $d$  monotonically increases. This is typical of both conventional and de Vries smectics, although as discussed above, it is somewhat contrary to the diffuse cone picture. The  $T$  dependence of the layer spacing at the transition depends crucially on the size of the parameter  $g'_1$ . In systems where the coupling of the layering and orientational order depends weakly on layer spacing and  $|g'_1|$  is unusually small, the  $T$ -dependence of  $q$  is almost flat. We have shown that if  $g'_1 = 0$  then the change in layer spacing scales like  $(T_{AC} - T)^3$ , and hence varies very weakly in the  $C$  phase near the  $AC$  transition. Systems with larger values of  $|g'_1|$  will have conventional behavior of the layer spacing.

Since this de Vries behavior of  $q$  has a different criterion than the de Vries behavior of  $M$  (which relies on  $t_n$  being athermal), it should be possible to find systems which exhibit de Vries behavior of the birefringence, but *not* de Vries behavior of the layer spacing.

For systems that approach the  $C$  phase from the  $A'$  phase, along path (iii), the birefringence will increase monotonically before jumping substantially (on the order of  $M_0$ ) at the transition. From Eq. (13), we see that the  $T$  dependent piece of  $q'_A$  is second order in the coupling  $g_{10}$  which we treat perturbatively in our analysis. Thus, this  $T$  dependent piece is very small in the  $A$  phase. Upon entry into the  $C$  phase the  $T$  dependence of the layering spacing will be weak if  $g'_1$  is small. At the  $A'C$  transition there will be a jump in  $q$  but Eqs. (8), (11) and (13) can be used to show that this jump will be small when  $g'_1$  and  $g_2$  are small. For such a system a transition just above the CEP will exhibit an almost continuous change in tilt angle, a weakly varying layer spacing, a substantial jump in birefringence, and a latent heat. Systems entering the  $C$  phase well above this CEP will exhibit jumps in all of the above quantities and thus not be de Vries-like.

The requirement of near  $T$ -independence of  $t_n$  for de Vries behavior severely restricts the possible experimental loci in Fig. 1 that can display such behavior: namely, nearly horizontal ones. A significantly sloped path like (i) will *not* exhibit de Vries behavior. In this case the

growth of the “bare” (i.e., coupling-free) birefringence  $M_0(t_n)$  as  $T$  is lowered swamps the effects due to the coupling terms, and makes the behavior of both the birefringence and the layer spacing conventional. Thus, our model can accommodate *either* conventional behavior or de Vries behavior, if  $t_n$  is thermal or athermal, respectively.

We acknowledge very helpful discussions with Yuri Panarin. K.S., D.H, S.P. were sponsored by the Department of the Navy, Office of Naval Research.

\* Corresponding author: ksaunders@calpoly.edu

- 
- [1] For an excellent review of de Vries smectics, see J. P. F. Lagerwall and F. Giesselmann, *ChemPhysChem*, **7**, 20, (2006).
  - [2] J. P. F. Lagerwall, F. Giesselmann, M. D. Radcliffe, *Phys. Rev. E* **66**, 031703 (2002).
  - [3] Y. Panarin (private communication).
  - [4] N. Hayashi, A. Kocot, M.J. Linehan, A. Fukuda, J.K. Vij, G. Heppke, J. Naciri, S. Kawada and S. Kondoh, *Phys. Rev. E* **74**, 051706 (2006).
  - [5] A. de Vries, *Mol. Cryst. Liq. Cryst.* **41**, 27 (1977); **49**, 179, (1979).
  - [6] Unfortunately, published data on birefringence of these materials within the  $A$  phase seems only to focus very near the  $AC$  transition, although such nonmonotonicity has been seen in DS13-MR11 [3].
  - [7] P. Barois, J. Prost, and T. C. Lubensky, *J. Phys. (Paris)* **46**, 391 (1985).
  - [8] G. Grinstein and R. A. Pelcovits, *Phys. Rev. A* **26**, 2196 (1982).
  - [9] S. A. Brazovskii and S. G. Dmitriev, *Zh. Eksp. Teor. Fiz.* **69**, 979 (1975) [*Sov. Phys. JETP* **42**, 497 (1976)].
  - [10] K. Saunders, J. Toner (in preparation).
  - [11] P. G. De Gennes and J. Prost, *The Physics of Liquid Crystals* (Oxford University Press, New York, 1995).
  - [12] It should be noted that for the system to exhibit an  $NAC$  multicritical point a term  $(q_i q_j Q_{ij})^2 |\Psi|^2$  must be included.
  - [13] We have for simplicity dropped some terms (e.g.,  $Tr(Q^2)|\Psi|^2$ ) that are actually lower order in the presumed small fields  $\Psi$  and  $Q$  than the terms we have kept. We have verified that keeping such terms with small, but non-zero, coefficients has no qualitative effect on our results. In contrast, dropping any one of the terms we have kept is unphysical.
  - [14] A free energy similar to ours was used by P.K. Mukherjee et al (*J. Chem. Phys.* **117**, 7788, (2003)) to investigate first order  $IA$  and  $IC$  transitions. However, they ignored the biaxiality of the orientational order parameter, and incorrectly claimed that the  $C$  phase is stabilized by the  $g_2$  term. This last mistake led them to exclude the  $h$  term, which is necessary to stabilize the  $C$  phase.
  - [15] See for example, P.M. Chaikin and T.C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995).
  - [16] See EPAPS Document No. [ ] for equations for the phase boundaries. For more information on EPAPS, see <http://www.aip.org/pubservs/epaps.html>.

## **Executive Summary**

The goal of this project was to develop and characterize the process steps required to fabricate a micro-actuator structure based on NiTi shape-memory alloy thin films. Micro-actuators are a key component of many micro-systems, which are capable of detecting bio-hazardous agents, dispensing drugs to combat cancer or improving the success rate of inter-vascular surgery. Moreover, these sensors involve molecular and atomic interactions which can enable students to explore design techniques that bridge into the nanotechnology world. The roadmap for the project consisted of three phases (1-Design, 2-Test & 3-Implementation); the first two phases were completed under Project#53910 and the third phase was completed under Project#53956. Overall, the project completed all of the planned tasks within the assigned budgets and the results have been presented at both MRS and ASEE technical society meetings. In addition, a strong industrial partnership has been established with Olympus Micro Systems and a summer internship at their research facility in Atlanta, GA was completed during the summer of 2006.

### **Phase 1 – Design**

#### **Develop a process for sputter deposition of shape-memory NiTi thin films on silicon wafers.**

Utilizing a RF&DC powered sputtering system with high temperature vacuum annealing capabilities a process sequence was developed for depositing NiTi on silicon using chromium as an adhesion layer and encapsulant to prevent oxidation of the Ti on the surface. The process was designed to deposit atomically equal NiTi thin films on <100> silicon wafers by co-sputtering from NiTi and Ti targets at 250 and 75 watts of DC power, respectively. These power settings yielded a deposition rate of 12 angstroms per second and the film's stoichiometry was confirmed by SEM-EDS. Pre-sputtering the NiTi and Ti targets effectively cleaned their surfaces and reduced the sputtering systems base pressure to  $10^{-7}$  torr thereby minimized the oxygen content of the NiTi films, which is crucial for shape-memory behavior. It was determined that a 100 nm chromium layer provided excellent adhesion between NiTi and silicon layers. The optimum annealing temperature to form a NiTi crystalline structure that would demonstrate super elastic behavior was verified through x-ray diffraction measurements. A fundamental understanding of the relationship of crystalline structure, film stoichiometry and phase transformation temperatures were established, which are necessary for designing a cantilever style micro-actuator.

### **Phase 2 – Test**

#### **Characterize the shape-memory behavior, electrical properties and stresses in NiTi films.**

The force displacement curves for both the martensite and austenite phases, that are characteristic of the shape-memory transition, were measured using a profilometer. From this data, the elastic modulus of the two phases of the NiTi film can be determined. Next the stresses in NiTi and Cr thin films were characterized as a function of sputtering power and pressure. Then the electrical resistance of the two NiTi phases was determined using a 4-point probe. The results from these experiments will then enable the design of a cantilever actuator where the residual stresses in the film stack are balance against the changes in the elastic modulus of the NiTi as it is thermally cycled through its phase

transformation. The balance of forces will result in a net displacement of the cantilever tip as it is ohmically heated through the NiTi phase transformation temperature.

### **Phase 3 – Implementation**

#### **Formulate a process for fabricating a cantilever style micro-actuator.**

The final step in the project is to develop a microfabrication process for forming free-floating cantilevers on a silicon substrate. Process steps were developed to deposit each material layer (e.g. silicon dioxide, chromium, NiTi, chromium) and then pattern the layers using photolithography to form the cantilever device. Each layer must be etched down to the silicon and then the silicon must be etched underneath the cantilever to release it and enable the formation of an actuator. Etching the silicon to release the cantilever is the most difficult part of this process and wet KOH process was characterized and optimized for this purpose.

#### **Broader Impact**

A Micro Systems Technology (MST) group has been formed at Cal Poly which includes a diverse group of undergraduate and graduate students from the Materials Engineering, Electrical Engineering, Industrial Manufacturing Engineering and Biomedical Engineering departments. By equipping undergraduate engineering students with the cross-disciplinary tools required to complete this project, they have been enabled to participate in the emerging micro and nano technology based industries. As students practice engineering in the coming decades, they must be equipped to: 1) understand the impact of scaling on material properties and engineering principles, 2) approach engineering from a systems perspective which emphasizes holistic thinking, 3) understand the interdisciplinary nature of micro & nano technologies and 4) demonstrate self-directed learning.

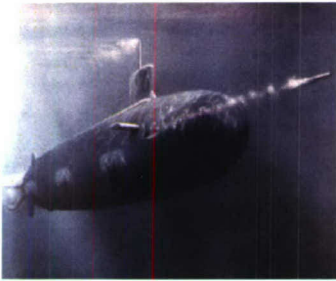
Results from this work have been presented at a Materials Research Society conference<sup>1</sup> and in part at the recent ASEE Pacific Southwest conference on Nanotechnology & Preparing Engineers for 2020<sup>2</sup>. This project has supported three masters' thesis students and five undergraduate senior projects.<sup>4-11</sup> In addition, a proposal has been submitted to the NSF<sup>3</sup> for a CCLI grant to develop a graduate-level MST course that will include some of the microfabrication techniques developed during this project. The MST micorfab lab occupies a newly constructed 1,500 ft<sup>2</sup> Class 1000 clean room facility which houses a basic set of thin film processing equipment for depositing, patterning and etching thin films as well as characterizing their behavior. This lab serves as the support facility for our current microfabrication laboratory course and for a new course being jointly developed with the Mechanical Engineering department based on the design and fabrication of a microfluidic device.

#### **Relevance of the Project**

##### ***The Technology of Micro Systems***

These tiny machines are too small to be seen by the unaided eye but are having a huge impact on the way we live, work and play.<sup>12</sup> One of the first commercial applications of micro systems or micro electro mechanical systems (MEMS) was the airbag. It relies on a micro accelerometer to sense a car crash and signal the deployment of the airbag.

This technology also enabled the fabrication of thousands of micro-mirrors, which produce digital light processing (DLP) arrays; DLP chips are found in most all tabletop PC-projectors today. Recently, IBM has announced that ThinkPad® laptops will contain an active protection system (APS) that detects if the computer is dropped and safely parks the hard drive to minimize data loss; this APS technology is also based on a MEMS accelerometer.<sup>13</sup> Inkjet printer heads contain MEMS devices and RF-MEMS devices play a key role in mobile phones, global positioning systems and smart missile targeting systems. Micro systems are found throughout many defense systems, which use MEMS-based digital inertial guidance devices, such as the Predator Missile program, shoulder fired anti-tank weapons and “smart” bombs.



(Note: Figure(s) are from ref.12)

Examples of micro systems are also spreading throughout the medical and health care fields. Pacemakers, retinal implants, hearing aids, blood pressure & glucose monitors for diabetics that require 1/3 the blood sample size as conventional macro instruments, all have MEMS components. Imagine machines so tiny that they can flow through the bloodstream, targeting cancer cells for elimination; you have now entered the nanotechnology world.

Multiple diseases can be screened with lab-on-a-chip devices that employ micro fluidic channels that are narrower than a human hair, but can quickly separate liquids and gases and allow micro sensors to analyze their properties. These micro sensors utilize thousands of cantilevers (imagine a device shaped like a small diving board) fabricated into detection arrays. Each cantilever is coated with a bio-film that selectively bonds to certain molecules that are associated with specific diseases. The cantilevers bend when the molecules are adsorbed onto their surface, producing a measurable change in their capacity to store electric charge, which can be measured by built-in electronic circuits. The entire device can be mass produced like memory computer chips and sold for a few dollars each making advanced diagnostics available to everyone at a reasonable cost.

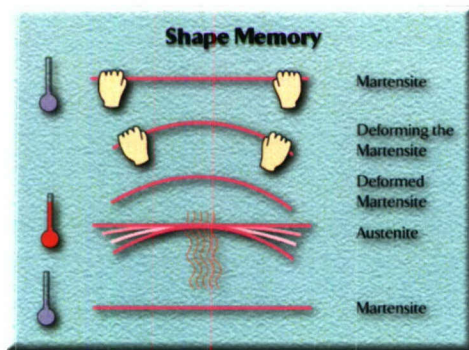
Geneticists use these devices to conduct DNA and protein analysis on a chip the size of a dime and get results in 30 minutes, instead of days or weeks from a commercial lab. Soldiers will be able to detect airborne pathogens, such as chemical and biological weapons and instantly identify the threat before it reaches lethal concentrations.

#### ***NiTi Shape Memory Alloy Thin Films***

Shape memory alloys (SMA) are metals that exhibit two very unique properties: pseudo-elasticity and the shaped memory effect. Ame Olander first observed these unusual properties in 1938 and they have been researched widely for use in many medical

applications including vascular stents and endovascular microtools.<sup>14</sup> An endovascular micro device consists of an active bending catheter (using SMA materials) and a micro drug infusion device for in-vivo patient testing and treatment.

The process for forming actuators made of SMA metals is as follows: 1) sputter deposit a 50/50 atomic weight percent NiTi thin film on silicon (it will be amorphous or non-crystalline), 2) anneal at a high temperature to form the austenite crystalline structure (rigid), 3) cool the alloy to room temperature and it will undergo a solid-state phase transformation to the martensite crystalline structure (which is relatively soft and ductile), 4) residual stress in the underlying poly-silicon layer will cause the film to deform and bend at room temperature, then 5) heat the film back up to about 100 °C and it will return to the original shape it had when annealed. This cycle can be used to produce actuators.



Actuators using these films demonstrate large force/mass ratios and induce high spatial displacements. This enables them to produce the highest work output of any actuator.<sup>15</sup> NiTi SMA micro grippers have been fabricated that are capable of 8% recoverable strain and 500 MPa of recoverable stress; they have demonstrated a work density as high as  $5 \times 10^7$  J/M.<sup>16</sup>

### ***Applications to Nanotechnology***

Nanotechnology involves materials at the nanometer level (1 billionth of a meter); imagine the thickness of a redwood tree compared to a human hair; that's the difference between a human hair and something one nanometer in size. A good concise definition of nanotechnology would be the creation and application of structures that are 1-100 nanometers or nm ( $10^{-9}$  meters) in size. This includes viruses (25nm), hemoglobin (7nm) carbon nanotubes (1.4nm) and a single strand of DNA (2.5nm); in fact, atoms are on the scale of 0.1 nm in diameter. It is beyond the scope of this project proposal to try to discuss all of the applications that fall under the term nanotechnology, but we will focus on some medical areas that intersect with micro systems.

Micro systems can be the platforms that enable scientists to explore biological systems at the molecular level. The study of self-replicating molecular machines is revolutionizing our approach to the detection and dispensing of biological agents. A micro fluidic pump in a MEMS device could deliver dendrimers (tree-like polymers similar to nylon) with cell hunting capabilities that would target and deliver chemical treatments at the molecular level. At 2.5nm, a strand of DNA is about the same size as a carbon nanotube; biochemists

may want to take advantage of the similarity of scale and attach therapeutic molecules onto the carrier nanotubes and “program” them to interact with specific cells (e.g. cancerous cells).

### **Detailed Report for Phase 3 – Implementation**

Silicon dioxide diaphragms are a vital component to many microelectromechanical systems (MEMS) devices, such as microfluidic dispensers and pressure sensors. These devices must be micromachined using anisotropic wet etches. Anisotropic etches preferential etch in certain directions, due to the highly ordered silicon, single crystal substrate. Certain planes are connected with weaker bonds that are more readily affected by the etch. Resulting in etched trenches with defined angles walls along the (111) plane.

The two most widely used anisotropic wet etches are ethylene diamine-pyrocatechol (EDP) and potassium hydroxide (KOH). These commonly used etches have known etching properties but must be characterized for each etching process. Many parameters can affect the experimental etch rate including: etch concentration, temperature, and feature orientation on the surface of the silicon substrate.<sup>17</sup> EDP has previously been used in our labs for wet etching NiTi cantilevers. It was preferred over KOH for its compatibility with NiTi the desired mask material. EDP's selectivity in the (100) to (111) plane is 35:1. Making it well suited for micromachining free floating cantilevers in that it etches laterally but etches more selectively down in the (100) plane. However, EDP also presents a health and safety hazard to the students using it. Ethylenediamine causes allergic respiratory sensitization of the lungs and pyrocatechol is classified as toxic and corrosive according to their MSDS. When EDP etch was used in the lab, necessary safety precautions included wearing a ventilator and limiting the amount of students in the lab.

Potassium Hydroxide (KOH) was studied and characterized for the application of SiO<sub>2</sub> diaphragms due to its selectivity of the (100) to the (111) plane of 400:1.<sup>18</sup> Such a high selectivity allows for deep holes with high aspect ratios to be etched into the back of a silicon wafer, to depths around 450 microns. KOH also provides less of a safety risk to the students using it but students must be cautious to avoid contact with their eyes as it could cause blindness. Moreover, the biggest drawback to using KOH as a wet etch is its limited compatibility with different mask materials. Although for micromachining silicon diaphragms the most common mask materials are silicon dioxide or silicon nitride of which are both compatible with KOH.<sup>19</sup>

### ***Experimental Methods***

The goal of this project was to quantify the etch characteristics of KOH on a (100) silicon wafer with a silicon dioxide mask. The variables in the experiment were the concentration of the KOH etch bath at 2.5 and 5.0 molar and the temperature of the bath at 70°C and 75°C. Both the etch rate and the selectivity of etch silicon in proportion to the mask material were determined. All measurements were made with a stylus profilometer. And the amount of mask under-etching that occurred relative to the features orientation on the wafer were measured with an optical microscope.

### ***Sample Preparation***

Silicon wafer of (100) orientation were cleaned and placed in a horizontal oxidation furnace. The wafers were in the furnace for forty-eight minutes at 1040 °C while water vapor was allowed to flow into the chamber. The resulting oxide thickness when the wafers were removed from the furnace was approximately 0.48 μm. Oxide thickness

measurements were taken with an optical reflectometer. (Note: Oxide thickness varied from wafer to wafer due to the temperature gradients in the furnace. This was accounted for when calculating the etch rate and selectivity.)

The masked wafers were then patterned using photolithography. Each wafer was placed on a hot plate at 150 °C for two minutes to remove any water vapor that might be absorbed on the surface of the oxide.<sup>20</sup> The wafer was then transferred into the spin coater and primed with 2 mL of hexamethyldisilazane (HMDS); this forces the surface to become hydrophobic for a short time to promote good photoresist adhesion. The HMDS was spun off at 400 rpm for 5 sec. Then 3mL of positive photo resist was dispensed at 1000 rpm for 5 sec. The spin coater was then ramped up to 3000 rpm for 30 sec to reduce the amount of residual solvents in the photoresist. Upon removing the wafer from the spin coater it was post-baked for one minute at 90°C to further remove any remaining solvents. Now that the wafers were coated in photoresist they were ready to be patterned. The proximity aligner employs a mercury vapor lamp which was used to expose each of the wafers. But in order to determine the correct light integral needed to expose the mask pattern an exposure matrix was completed. The mask was loaded into the aligner and a polypropylene exposure plate, was placed over the mask allowing only a portion of the wafer to be exposed at a time. Each eighth of the wafer was exposed to a different light integrals ranging from 5.0 to 12.0. The wafer was then examined to determine where the light had fully penetrated the photoresist and where the resolution looked the best. The highest resolution corresponded to the region of the wafer that was exposed to a light integral of 7.0 but it was apparent that increasing the light integral would optimize the exposure so the remainder of the wafers were exposed to a 7.5 light integral at a print gap of 3µm.

The sample wafers were then loaded in to the aligner and were processed as a batch. After exposure they were developed in a 1:1 solution of positive photoresist developer and DI water for a minute and a half, and quenched in DI water. The wafers were then hard baked for one minute at 150 °C to improve the adhesion of the patterned photoresist to the oxide. A wafer with the same wet oxide layer was then step etched in a buffered oxide etch (BOE) [HF(49%):NH<sub>4</sub>F] in a 6:1 mixture at room temperature to determine the etch rate of the BOE on the wet oxide mask. Again using the reflectometer the thickness of the multiple step etched regions were plotted and the etch rate was determined to be 1.4 nm/sec. The patterned photoresist wafers were then allowed sit in the BOE etch for six minutes to assure the oxide mask was fully patterned. The photoresist layer was then removed in 1165 Microposit remover bath at 70 °C for 10 minutes.

### ***Wet Etching***

Before any wet etching occurred into the silicon substrate, initial oxide thickness measurements were taken with a XP-1 Stylus Profiler made by Ambious Technologies. The profilometer was capable of measuring precision step heights from under 10 angstroms to as large as 100 micron.<sup>21</sup> The profilometer moved a cantilever tip of 2.5 microns in diameter across a 1 mm section of the substrate at a speed of 0.1 mm/sec. This type of scan was repeated ten times on each sample to determine the average oxide step height thickness. The etch baths either 2.5 mol or 5.0 mol concentration. The KOH as received form was pellets that were weighted and then added to DI water to make a liter of solution. The etches were then heated to their desired temperature plus five degrees that accounted

for the heat transfer between the solution and the cassette when it was lowered into the etch bath. The temperature was constantly monitored for the duration of the etch time to assure no fluctuations occurred.

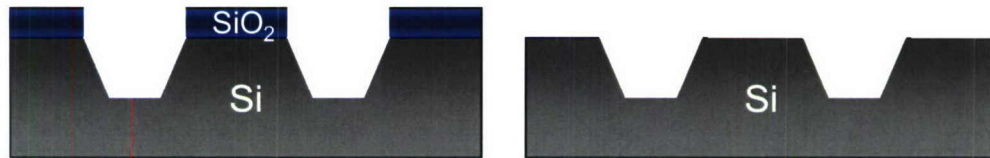
Wet Etching procedures for each wafer:

1. The oxide step height was measured with 10 profilometer scans
2. The wafer was placed in the KOH solution for five minutes, was DI quenched and Nitrogen dried.
3. Ten step height scans were complete to determine the depth of the etched region.
4. Steps 2 and 3 were repeated at 10 min, 15 min, 20 min, 40 min, and 60 min.

The averages of each of the etch intervals were plotted to determine the etch rate of KOH at the experiments given parameters.

### ***Selectivity $\text{SiO}_2\text{:Si}$ Measurements***

A portion of the etch wafer was cleaved off along the [100] direction. Step profiles of the surface of the wafer were taken to determine the depth of the etch region of the sample. The sample was then loaded in to a single wafer holder and lowered in to the buffered oxide etch (BOE) to remove the remaining oxide layer. The step height was measured once again to determine the depth of the silicon etched (Figure 1).



**Figure 1.** Cross-sectional view of etched silicon wafer with and without oxide mask.

The measurement averages of the data before the BOE and after were compared to determine the remaining oxide thickness after the etch process. The selectivity calculations are shown below.

$$\frac{O_i - O_f}{|S_i - S_f|} = \frac{O_e}{S_f} \quad (\text{Eq. 1})$$

$O_i$  = Initial oxide thickness

$O_f$  = Final oxide thickness

$O_e$  = Oxide Etched

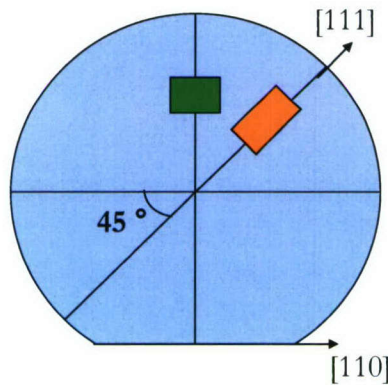
$S_i$  = Initial Silicon etched

$S_f$  = Final Silicon etched

$$\left[ \frac{\left( \frac{O_e}{S_f} \right)}{\text{MinutesEtched}} \right] = \text{selectivity} \quad (\text{Eq. 2})$$

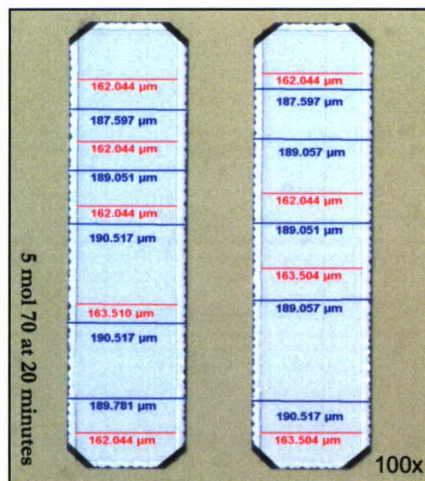
### Mask Underetching Measurements

Due to the ordered crystallographic nature of the silicon substrate, and the preferential etching behavior of KOH in certain directions. Underetching only occurs when the feature being etch is turned 45 degrees relative to the major flat (Figure 2).



**Figure 2.** This is a top view of a (100) Si wafer. The green box represents region that is parallel to the major flat and the [110] direction, in which no underetching will occur. The orange box represents a feature that is parallel to the [111] direction which will experience underetching.

To measure the amount of underetching an Olympus optical microscope was calibrated and used to measure the widths of the mask opening and the width of the underlining etched region of a feature turned 45 degrees to the major flat (Figure 3). The widths were compared to determine the percent mask underetching that occurred.



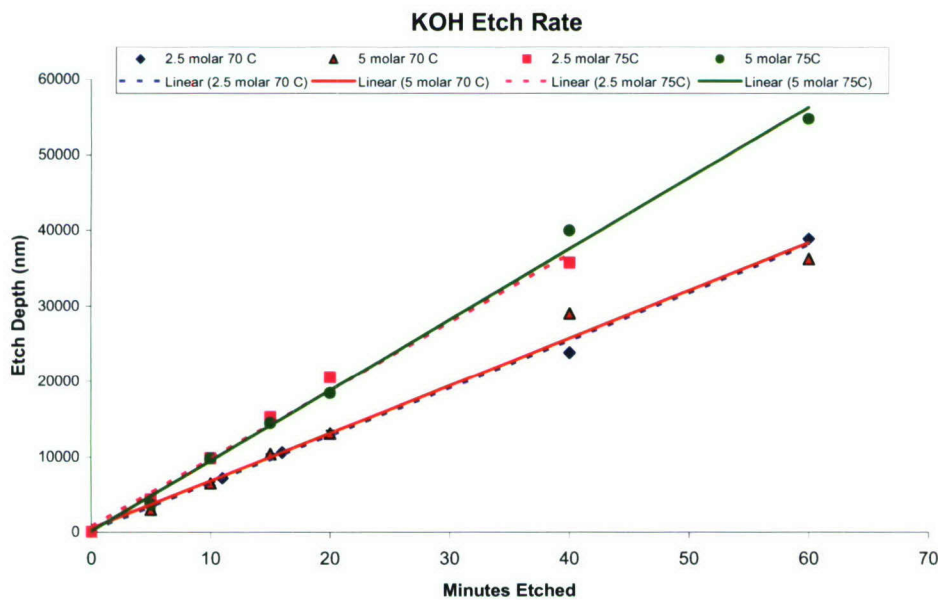
**Figure 3.** An etched feature parallel to the [111] direction on a (100) Si wafer. The red lines are width measurements of the mask opening and the blue line are the width of the underlining etched region.

### Results

The following sections summarize the KOH etch rate for each of the experimental parameters, as well as, the mask to substrate selectivity and the amount of underetching.

### Quantifying KOH Etch Rate

The data collected on the etch depths of each of the sample is presented in Figure 4. The dashed lines in the graph represent the wafers that were etched in 2.5 molar solutions, and the solid lines represent the 5 molar solution. Error bars representing one standard deviation are present but are not visible due to the consistency of the data. The calculated etch rate from the slope of each of the lines is presented in Table I.



**Figure 4.** Etch data for concentration vs. temperature study. The slope of the given trend lines represent the calculated etch rate of the solution.

**Table I.** Experimental KOH Etch Rate of Si Wafer

Experimental Parameter	Etch Rate	Error (nm)
70°C 2.5 mol	630.5 nm/min	± 70
70°C 5.0 mol	631.7 nm/min	± 17
75°C 2.5 mol	902.6 nm/min	± 206
75°C 5.0 mol	936.0 nm/min	± 62

### Quantifying Selectivity of SiO<sub>2</sub>:Si

The selectivity measurements of the SiO<sub>2</sub> mask to the Si substrate are presented in Table II. The data was reduced to a ratio for quick reference.

**Table II.** Selectivity measurements of SiO<sub>2</sub>:Si

Experimental Parameter	Selectivity SiO <sub>2</sub> :Si	Selectivity Ratio
70°C 2.5 mol	3.8 : 589 nm/min	1 : 148
70°C 5.0 mol	3.4 : 660 nm/min	1 : 194
75°C 2.5 mol	3.5 : 900 nm/min	1 : 257
75°C 5.0 mol	4.0 : 915 nm/min	1 : 229

### ***Underetching***

This measurement was only preformed on two samples: 5 molar at 70°C for 20 minutes and 2.5 molar at 70°C for 60 minutes. The feature examined was parallel the [111] direction. The data collected is presented in Table III.

**Table III.** Percent underetching that occurred on structures parallel to the [111] direction.

Experimental Parameters	Minutes Etched	Percent Etched
70°C 2.5 mol	20	13 %
70°C 5.0 mol	60	33 %

### ***Discussion***

#### ***KOH etch rate***

Upon completion of the etch experiments it was obvious that the KOH etch rate was dependant on temperature and not concentration. The two samples that were etched in 70°C solutions experienced etch rates around 631 nm/min. Where as the samples in solutions that were only five degrees higher experienced etch rates nearly 300 nanometers a minute more. When the concentration of the solutions were varied the etch rate did not change significantly. Initially it was thought that for longer etch times the etch rate of the lower concentration solutions would diminish as the reactive components were used up. But for the time scale tested no such event occurred. When KOH is later implemented in the lab setting, lower concentration solutions can be used and less KOH pellets will be needed for the overall experiment. At lower concentrations the KOH etch reaction is diffusion dependant so it will be important to agitate the etch bath to attain consistent etch results.<sup>21</sup>

#### ***Selectivity of SiO<sub>2</sub>:Si***

The selectivity results were consistent with the etch rate data. Again there was a trend with temperature but no significant difference when the concentration was varied. The experimental results differed from what was expected in literature. The differences could be attributed to the type of oxide mask, the lack of agitation during the etch process, and the back surface of the wafer was not protected during the etch process. The selectivity of the mask to the wafer was highly dependant on the type of oxide mask grown on the surface of the silicon wafer. As mentioned previously the oxide mask was wet thermally grown resulting in a lower density oxide mask than if it had be deposited with plasma. The less dense structure of the oxide would have had a higher concentration of hydrogen ions in the lattice. When the hydrogen ions came in contact with the KOH they would have acted as a surface barrier on the oxide mask and kept the KOH from etching it. Producing better selectivity for SiO<sub>2</sub>:Si, than the value predicted in empirical calculation of selectivity. If the etch bath was agitated the selectivity of the oxide to the silicon wafer would greatly decrease

From the etch measurements it was easy to see that much of the silicon substrate was being etched but since the back side of the wafer was not protected. This could have affected the etch rate data and also the selectivity measurements if the stylus of the profilometer was not in the same relative height each time due to the reduction of the wafer thickness. But the stylus on the profilometer makes measurements relative to the top of the sample, so it should not matter that there is uniform etching occurring on the back side of the wafer.

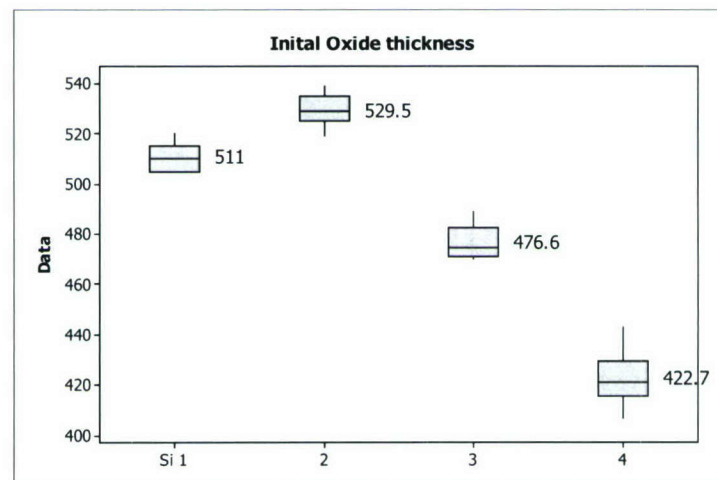
### ***Underetching***

As described before underetching on (100) Si wafers with anisotropic etches such as KOH attack the substrate preferentially. That is, they etch at a higher rate in certain directions than they do in others, exposing the slower etching planes over time. Since KOH is said to have a selectivity of 400:1 ((100):(111)) for the application of making silicon dioxide diagrams it is preferred to limit the amount of mask underetching.<sup>18</sup>

When a mask is later made for the processing of silicon dioxide diaphragms it would be beneficial to place the structures parallel to the major flat. Structures measured with this orientation of the wafer did not experience any underetching and retained dimensional integrity of the initial patterned feature.

### ***Experimental Error and Process Challenges***

The initial oxide thickness of each of the samples varied due to their relative position in the furnace, during the wet oxidation growth step. Figure 5 shows the initial thickness of the oxide for each of the sample wafers.



**Figure 5.** Average oxide thickness of the sample wafers

During the photolithography stage of the experiment: The initial two wafers (1 and 2 in Figure 5) had a slight delamination problem with the photoresist, where the borders of the features began to lift off affecting the dimensional integrity of the test regions on the wafer. This occurred especially on the half of the wafer that had the larger features. The later wafers in the batch experienced even more delamination where approximately 70% of the wafers were not useful for collecting data. Areas where large portions of the photoresist sheeted off during developing could have been attributed to inconsistent timing between loading the spin coater and dispensing of HMDS and photoresist. In which a thin layer of

water vapor could have been permitted to form. So, when the wafer was placed in the developer the photoresist was able to detach from the surface of the wafer in sheets.

### ***Future Work***

This project still has room to grow. Additional work could be done to further categorize KOH's effect on silicon dioxide at other temperatures and concentrations. In my work I over shot the optimum concentration of KOH which is about 22 wt %.<sup>20</sup> Also I did not use agitation in my etch experiments, which could further accelerate the etch rate but for SiO<sub>2</sub> mask it will reduce the selectivity. Beyond changing the variables of the etch solution these experiments should be carried out on a number of other potential mask materials. KOH should have about the same selectivity with a chrome mask, additional mask materials to try would be silicon nitride, chrome nickel-titanium sandwich, and gold.[3] These mask should all be compatible but should be characterized for our own lab use. Particularly for the application of silicon diaphragms it might be important to repeat these same experiments on a (110) oriented Si wafer. Due to KOH's high selectivity in the (110) plane shorter etch times may be used to achieve the same deep rectangular features in less time.<sup>22</sup>

Beyond improving the level of safety in the lab KOH can be used in many bulk micromachining applications. EDP the previously used anisotropic wet etch's byproducts were classified as hazardous waste. Where as KOH a strong base can be neutralized and reintroduced into the waste stream. Although for our lab the waste will be treated the same and will be removed by the environmental health and safety department. KOH is also much less expensive than EDP. It only required the purchase of one chemical and etch baths can be complete in regular Pyrex beakers. EDP on the other hand required a careful mixture of four different chemicals and must be preformed in an elaborate test environment.

### ***Conclusions***

KOH and anisotropic etch was studied with a SiO<sub>2</sub> mask, it was determined that the etch rate was temperature not concentration dependant. When this data is utilized for the fabrication of silicon diaphragms, lower concentrations of etch baths can be used to reduce the cost of the experiment. The selectivity of the SiO<sub>2</sub> mask to the silicon substrate followed the same pattern, the selectivity increase with temperature but showed no correlation with an increase in concentration. The selectivity measures were inline with what was expected in literature<sup>23</sup> for the wet thermal oxide mask that was used. Adding and the variable of etch bath agitation would further optimize the etch rate but would reduce the selectivity. Due to the surface barrier of hydrogen bubbles that build up on the oxide would be removed and the oxide mask would more readily be etched. To limit the occurrence of underetching for making deep holes with high aspect ratios the feature should be oriented parallel to the major flat or the [110] direction. For features parallel to the [111] direction experienced 33% underetching after only one hour.

### **Acknowledgements**

The following students contributed to the majority of the work presented in this report: Audrey Medford, Brent Huigens, Dustin Dequine and David Getchel.

This work was sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855

## References

1. D. Getchel & R.N. Savage, "Characterization of Structural and Mechanical Properties of NiTi Shape Memory Thin-films for Micro Actuators," MRS Spring Conference, April 2005
2. R.N. Savage, "Micro Systems for Exploring the Nanotechnology World," ASEE-PSW Conference 2005 – Vision of Engineering in Year 2020, Loyola Marymount University, Los Angeles, CA, April 2005
3. NSF Proposal# 0535629 for NSF05-559 CCLI-Phase1: Exploratory, "Micro Systems for Exploring the Nanotechnology World," submitted 5/17/05
4. R. Hosler, "Characterization of Sputter-Deposited NiTi Thin Films Capable of Powering MEMS Devices," Senior Project Report, Materials Engineering, Cal Poly State University, June 2004
5. G. Rom, "Design and Characterization of Sputter-Deposited NiTi Thin Films for MEMS Actuation," Senior Project Report, Materials Engineering, Cal Poly State University, June 2004
6. A. Adams, "Characterization of Polysilicon Thin Films Deposited by Physical Vapor Deposition," Senior Project Report, Materials Engineering, Cal Poly State University, June 2005
7. O. Acton, "Developing Micro System Actuators: Residual Stress and Composition in Co-Sputtered Amorphous Thin Films," Senior Project Report, Materials Engineering, Cal Poly State University, June 2005
8. A. Medford, "Characterization of KOH Anisotropic Silicon Wet Etching Process," Senior Project Report, Materials Engineering, Cal Poly State University, June 2006
9. D. Dequine, "Trends in Thin Film Stress as a Function of Deposition Power and Pressure," Masters Thesis, Materials Engineering, Cal Poly State University, anticipated Spring 2007.
10. B. Huigens, "Electrical Properties of Thin Film NiTi for MEMS Actuator Applications," anticipated Spring 2007.
11. D. Getchel, "Characterization of Structural and Mechanical Properties of NiTi Shape Memory Thin-films for Cantilever-style Micro Actuators," anticipated Spring 2007.
12. Small Tech 101 – An Introduction to Micro and Nanotechnology, SmallTimes Media Inc. [www.smalltimes.com/products](http://www.smalltimes.com/products)
13. SmallTimes: Big News in Small Tech, SmallTimes Media, LLC, Vol 3, No 7, Nov/Dec 2003
14. J. Chang, et.al., "Development of Endovascular Microtools", Journal of Micromechanics and Microengineering, 12, 2002, pgs 824-831
15. P. Krulevitch, et.al., "Thin Film Shape Memory Alloy Microactuators", Journal MEMS, Vol. 5, No.4, 1996, pgs 270-282
16. W. Huang, et.al. "Design, Testing and Simulation of NiTi Shape-memory-alloy Thin-film Based Microgrippers", J. Microlithography, Microfabrication and Microsystems, VI.2, No.3, July 2003, pgs 185-190
17. Hai Ni, Hoo-Jeong Lee and Ainissa G. Ramirez, "A robust two-step etching process for large-scale microfabricated SiO<sub>2</sub> and Si<sub>3</sub>N<sub>4</sub> MEMS membranes," Sensors and Actuators A: Physical, Volume 119, Issue 2, , 13 April 2005, Pages 553-558.
18. G. Kovacs, N. Maluf, K. Peterson, "Bulk Micromachining of Silicon," Proceedings of the IEEE, Volume 86, No. 8, August 1998, Pages 1536-1551.

19. Williams, "Etch Rates for Micromachining Processing – Part II," Journal of Microelectromechanical Systems, volume 12, No. 6, December 2003, Pages 761-778.
20. D. Sheeja, B. K. Tay, L. J. Yu, D. H. C. Chua, W. I. Milne, J. Miao and Y. Q. Fu, "Fabrication of amorphous carbon cantilever structures by isotropic and anisotropic wet etching methods," Diamond and Related Materials, Volume 12, Issue 9 , September 2003, Pages 1495-1499.
21. Ambious Technology. XP Series Stylus Profiler User's Manual (2003).
22. Rai-Choudhury, P. Handbook of Micromachining and Microfabrication. Vol. 2. IEE Materials and Devices 12B. Bellingham: SPIE Optical Engineering, 1997, Pages 57-65.
23. K. E. Bean, "Anisotropic Etching of Silicon," IEEE Trans. Electron Devices, vol. ED-25, October 1978, Pages 1185-1193.

**DNA-linked polystyrene microstructures: Selective DNA attachment and enhanced electrostatic stabilization**

Project Investigator:

Peter V. Schwartz  
Physics Department  
California Polytechnic State University  
San Luis Obispo, California

## 1. Abstract

Two DNA strands will bond to each other, forming a double strand of DNA, only if the sequences of the two strands are compatible. Therefore, DNA can be used as a sequence specific Velcro<sup>®</sup> to assemble structures on the nanometer scale. By selectively attaching DNA to nanostructures, we have shown DNA to be an excellent selective nanoscopic adhesive to guide the assembly of tailored materials. Funded by C<sup>3</sup>RP, Cal Poly undergraduate students have produced a working system whereby DNA-decorated microspheres assemble controllably, reproducibly, and reversibly. The work resulted in a peer reviewed publication spring of 2005.<sup>1</sup> This project is truly multidisciplinary, involving faculty on Cal Poly campus from physics, chemistry, biology, and material engineering. We continue our work with new and old collaborators at Cal Poly, the physics department of New York University (NYU), and in chemical engineering and Materials Research Laboratory at University of California at Santa Barbara (UCSB). At Cal Poly, we are optimizing the DNA functionalization process by varying the parameters during the attachment process. Are students are quantifying DNA surface density by fluorescence with a flow cytometer at UCSB. At NYU, Dave Pine heads a research team developing novel polystyrene building blocks. Two physics students and I visited NYU this summer to show Pine's group the DNA technology developed at Cal Poly, and help them adapt this technology to the novel structures being developed at NYU. Pine was very impressed with our students, finding remarkable that they were both sophomores. Additionally, working with Craig Hawker (Biochemistry UCSB) we are developing a new way to enhance the passivation of the polystyrene surfaces by constructing hydrophilic peptide hyperbranches on the polystyrene surface. We wish to prevent non-DNA binding of the surfaces, and enhance the movement of polystyrene structures past each other. The goal is an annealing process whereby the assembly of polystyrene microstructures assumes a lowest energy, highly ordered, structure. The resulting colloidal crystal promises many structural, electronic, and optical attributes including that of a full photonic band gap.

## 2. Significance

The production of a full optical band-gap photonic crystal is a significant achievement as a micro-optics technology including low threshold lasers, optical waveguides, and optical multiplexing.<sup>2</sup> Yet, possibly the greater achievement toward a general nanotechnology is the demonstration that self-assembly with DNA linkage<sup>3</sup> can be used to render a technologically relevant substance with a predetermined architecture. The assembly of large molecules as well

---

<sup>1</sup> "Selective, Controllable, and Reversible Aggregation of Polystyrene Latex Microspheres via DNA Hybridization" Rogers, P. H.; Michel, E.; Bauer, C. A.; Vanderet, S.; Hansen, D.; Roberts, B. K.; Calvez, A.; Crews, J. B.; Lau, K. O.; Wood, A.; Pine, D. J.; Schwartz, P. V., *Langmuir*, 2005; 21(12); 5562-5569

<sup>2</sup> "Photonic Crystals, Molding the Flow of Light", John D. Joannopoulos, Robert D. Meade, Joshua N. Winn, Princeton University Press, 1995

<sup>3</sup> "DNA Nicks and Nodes and Nanotechnology", N. C. Seeman, *Nano Letters*, 2001, 1, 22; "Nanotechnology and the Double Helix" Seeman, N. C. *Scientific American*, 2004, 290 (6), 64-75; "Meniscus Force Nanografting: Nanoscopic Patterning of DNA" Schwartz, P. V. *Langmuir*, 2001, 17, 5971-5977; "Direct Patterning of Modified Oligonucleotides on Metals and Insulators by Dip-Pen Nanolithography", Demers, L. M.; Ginger, D. S.; Li, Z.; Chung, S. -W.; Mirkin, C. A. *Science*, 2002, 296, 1836-1838; "One-pot colorimetric differentiation of polynucleotides with single base imperfections using gold nanoparticle probes", Storhoff, J.J., Elghanian, R., Mucic, R.C., Mirkin, C.A., Letsinger, R.L. *J. Am. Chem. Soc.* 1998, 120, 1959; Bashir, R.; Superlattices and microstructures 2001, 29(1), 1-16; "Joining and Scission in the Self Assembly of Nanotubes from DNA Tiles" A. Ekani-Nkodo, A. Kumar, D. Kuchnir Fyngenson, *Phys.Rev.Lett.* 2004, to appear.

as organic and inorganic nanostructures can be guided by a number of mechanisms besides DNA hybridization including hydrophobic interactions, covalent bonds, protein interactions, and sedimentation. Two attributes of DNA hybridization (our process) that make it preferable to these other methods are: the bonding is sequence specific, allowing for an almost inexhaustible number of independent interactions; and the bonding is predictably reversible with increased temperature, which potentially allows one to anneal defects from a DNA-linked crystal. We have recently developed the ability to selectively attach microspheres into large aggregates by means of DNA hybridization under nonequilibrium conditions. However, the success of assembly by DNA hybridization toward a reliable nanotechnology lies in the ability to anneal an assembly into an equilibrium structure. It is this annealing process which we propose to study through the assembly and annealing of sub micron polystyrene spheres and other structures of linear and tetrahedral symmetry.

Besides yielding results related to particle self assembly, DNA-guided self-assembly, and photonic band gap crystal formation, the study of surface interactions with biologically relevant materials will have general biomedical and biotechnological applications. In particular, surfaces that do not bind to biological molecules (immunological responses) are of great importance. Additionally, DNA sequence recognition is of tremendous interest from a biology and medical perspective. Not only will our research result in publications, but the findings will be the subject of patent disclosures.

### **3. Other Support**

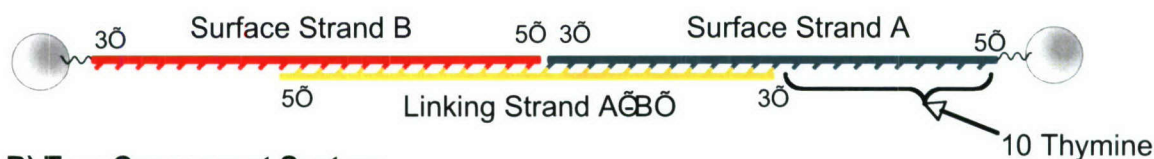
I have vigorously pursued external funding. This past three years, I have written several (5 or 6) research proposals both for projects that are related to the present proposal and other projects that are unrelated. A closely related proposal to Research Corporation, "Three-Dimensional Self-Assembly using DNA Linkage: Construction of Tetrahedral Photonic Bandgap Crystal" was rejected. A successful proposal to Petroleum Research Fund (PRF) entitled: "Molecular Transport from an Atomic Force Microscope Tip: The Basic Science Behind Dip-Pen Nanolithography" has a very small overlap with the present proposal. This \$50,000 three-year pending grant. The work of the present proposal has so far been funded by a \$20,000 grant from Raytheon, and from previous C<sup>3</sup>RP grants: \$20,000 in 2002, and \$30,000 in 2004 and 2005.

### **4. Detailed Plan of Work**

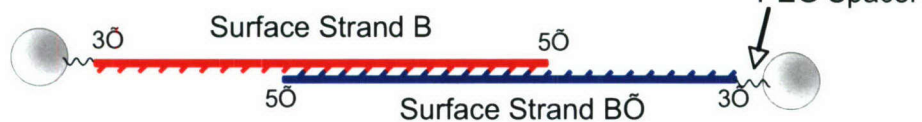
#### a) Past Work

Our study of the controlled aggregation of polystyrene beads (diameter = 1  $\mu\text{m}$ ) through hybridization of attached strands of DNA reached a successful first step with the publication of our first paper to *Langmuir*.<sup>1</sup> Besides myself, D. Pine, and E. Michel, the authors are nine present and past undergraduate students in physics and material engineering. David J. Pine was a professor at UCSB but has moved to NYU into a physics department specializing in materials research. Our collaboration will continue and will include summer employment for my research students in Pine's laboratory. Below is a summary of the 30 page manuscript, wherein we acknowledge: "Office of Naval Research, under Award # N00014-04-1-0436".

### A) Three-Component System

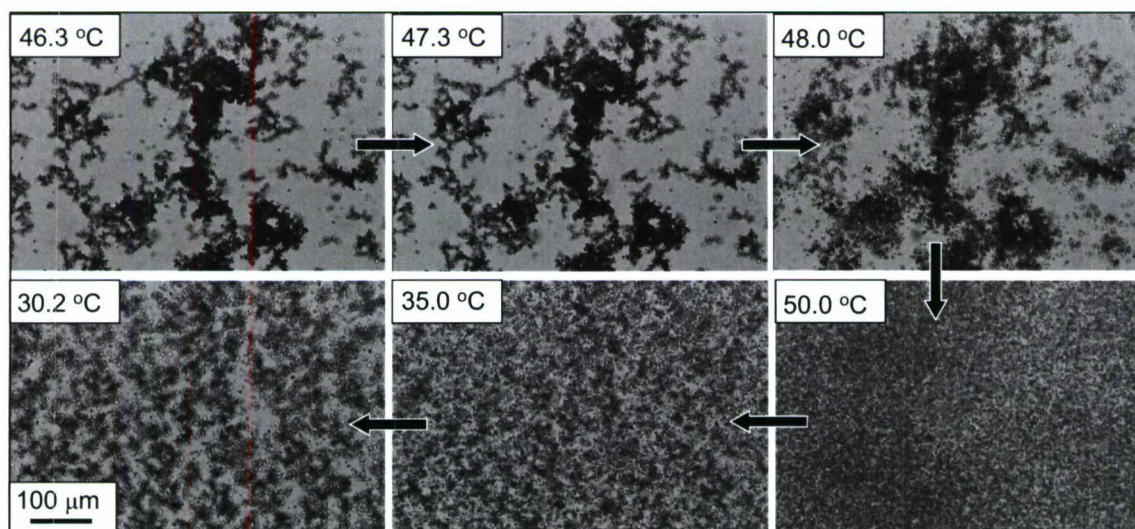


### B) Two-Component System



**Figure 1:** Attachment scheme for both the a), three-component system, and the b), two-component system

We developed a method to attach DNA to the surface of polystyrene microspheres at surface densities that are 20 to 50 times higher than accomplished by other researchers (corresponding to a molecular footprint of about  $5 \text{ nm}^2$ ). Using fluorescently-labeled DNA, we measured the surface density and the portion of attached DNA that hybridized to complementary DNA in solution (hybridization efficiency). DNA is negatively charged, and the resulting high surface charge was able to prevent the microspheres from clumping up – a common problem with microspheres. We were then able to demonstrate that the microspheres could be controllably and reversibly aggregated through hybridization between complementary strands of attached DNA on different microspheres (Fig. 1). We were able to reversibly thermally disperse aggregates by heating them through the melting temperature of the linking DNA (Fig. 2). Additionally, we used fluorescent DNA to show the selectivity of the DNA-guided assembly. Binding only occurred between microspheres that had complementary sequences of DNA.



**Figure 2:** dissociation and re-aggregation of a two component aggregate at 70 mM  $[\text{Na}^+]$ . Top Left: The beginning aggregate below dissociation temperature. Following the arrows, the sample is heated above the dissociation temperature ( $T_d = 47.3 \text{ }^\circ\text{C}$ ) and then allowed to cool and microspheres aggregate again.

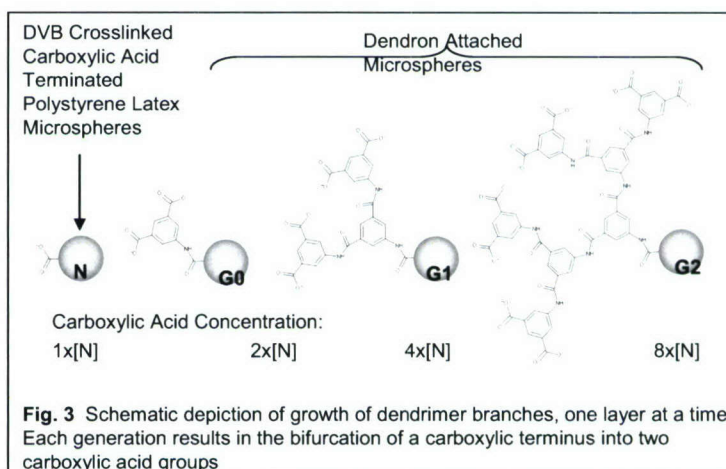
**Role of Surfactant in Nonspecific Binding.** It was found that the addition of Tween-20 at the end of the conjugation step enhances the ability of the aggregates to disperse upon heating. If Tween-20 is not used, a small core at the center of an aggregate may fail to disperse upon heating; including Tween-20 allows the complete thermal dissociation of the aggregate. We propose that the Tween-20 coated the hydrophobic patches on the polystyrene surfaces (making them hydrophilic), thus preventing hydrophobic-hydrophobic attraction between microspheres.

## b) Specific Design and Procedures

Our research goal consists of four parts: (i) Controlling nonspecific attraction without surfactants, (ii) Developing DNA systems that universally hybridize (iii) characterizing and optimizing the DNA attachment process (iv) building more complex crystals.

(i) Prevention of attraction between nonhybridized microstructures. It is imperative that if neighboring microspheres are not bound via DNA hybridization, they should be able to freely slide by each other in order to facilitate ordering during annealing. We hope to stabilize our particles without Tween by preventing short range hydrophobic attraction. We are been building hydrophilic hyperbranches with organic molecules as shown in Fig. 3. Our initial results were ambiguous. The organic molecules required us to conduct the experiments in an organic solvent, which swelled the polystyrene. We subsequently began new experiments using peptides for building branches because peptides are more hydrophilic and require no processing in troublesome organic solvents. These experiments have gone smoothly and the resulting microspheres are well stabilized. We will extend these experiments with DNA and with different lengths of attached polyethylene glycol (PEG), which stabilize microstructures with steric (entropic) repulsion.

(ii) Exploring microspheres which universally bond. In order to facilitate the movement of defect sites, a microsphere should be able to detach from one neighbor and bond to the next. This requires that every microsphere be able to bond with every other microsphere. As this is not possible with either of the hybridization systems illustrated in Fig. 1 (surface-bound strand A will not hybridize with itself), the annealing and melting behavior of “mongrel” microspheres which are coated with both varieties of surface-bound DNA will be studied along with a system utilizing a DNA sequence that binds to itself. This will result in a new variety of one-component, two-component and three-component hybridization schemes in which all microspheres in a batch can attach to every other microsphere. These new systems will be characterized with respect to melting and annealing behavior. The design and production of the new sequences of DNA has been facilitated by collaboration with Chad Immoos (Cal Poly chemistry) who has a DNA synthesizer and expertise in DNA synthesis.



(iii) Characterizing and optimizing the DNA attachment process. At Cal Poly, we have discovered that pH during the DNA conjugation (attachment) strongly affects the behavior of the resulting DNA-decorated microspheres. Across a wide pH, we will correlate conjugation pH with the behavior of the resulting microspheres, the DNA surface density, and the hybridization efficiency of the surface-bound DNA. DNA surface densities will be made with a flow cytometer at UCSB.

(iv) Building more complex crystals. David Pine of New York University produces the microspheres and has developed the ability to assemble microstructures into a controllable geometry, such as the tetramers shown in figure 4. . An example of the process whereby tetramers are assembled into a crystal of tetrahedral symmetry is shown in figure 5. In order to attach the DNA controllably to points on the microstructures, these tetramers must first be further processed to produce a “patchy particle”, shown in figure 6. The small protrusions in the spherical surface have a different chemical functionality than the rest of the polystyrene microstructure. The DNA selectively attaches only to these small protrusions. At the end of July, two students (Amber Bowen and Christopher Seck, Cal Poly physics majors) and I traveled to NYU to teach the Pine group the DNA technologies we developed at Cal Poly, and adapt these technologies to the unique microstructures Pine’s group is producing.

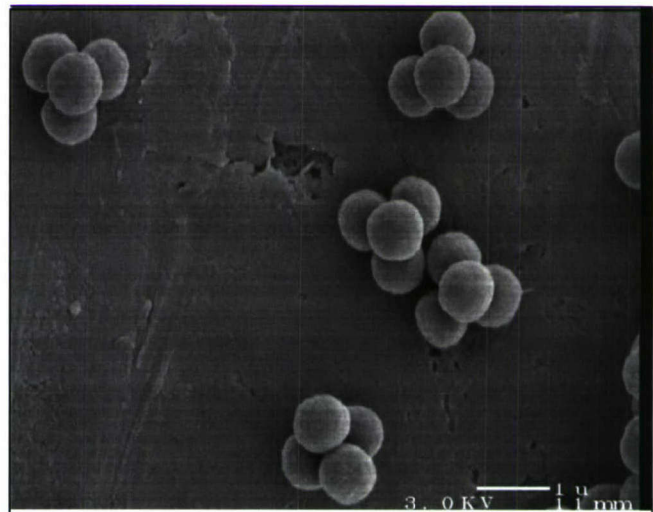
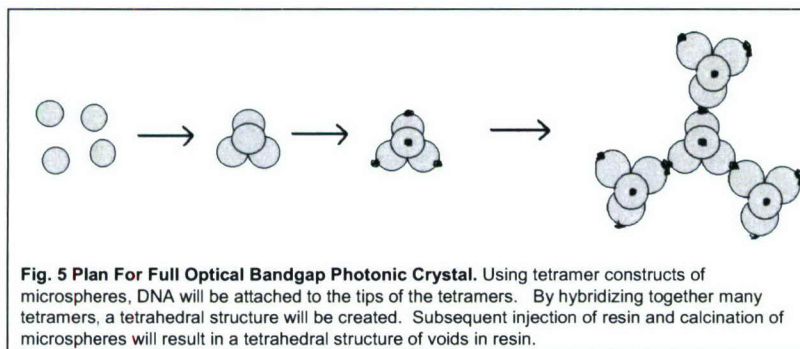


Fig. 4 tetramers. Scanning Electron Micrograph (SEM) of tetrahedral structures of microspheres (tetramers).



## Publications and Presentations Related to Project

### Abstracts

Seck, A. Bowen, M. Fischer, A. Leave, G. Tucker, B. Brasseur, B. Anthony, P. V. Schwartz; 'Selective, controllable, and reversible aggregation of polystyrene latex microspheres via DNA hybridization'; Presentation at COSAM, 2006.

By attaching short strands of DNA to polystyrene microspheres and allowing these beads to interact in solution with other beads that have complementary strands of DNA attached to them, we have been able to assemble aggregates of microspheres. We have also seen that the aggregates disperse at the melting temperature of the surface DNA, and the aggregates reform after allowing the dispersed solution to cool to room temperature. Lastly, we have been able to show that by adding ammonia, surfactants, salt concentrations and pH of the solution, we can vary the effects of nonspecific binding between the microspheres. This self-assembly of aggregates through DNA linkage could potentially be utilized in a process for making full photonic bandgap crystals. The key to creating full photonic bandgap crystals is purity; therefore, it is imperative that we are able to anneal the aggregates (in both the metallurgy and bio-physics sense of the term). We hope to show that the melting of aggregates will also show the same affects as annealing crystal structures. Once we can show that impurities can be corrected and controlled, we then hope to change the geometry of our particles, and take closer steps to creating a template for a tetrahedral photonic bandgap crystal with the assistance of dip pen nanolithography. Current work focuses on two specific areas. We are running a reaction in an organic solvent and subsequently attempting to deswell the microspheres to extract the organic solvent from the polystyrene matrix. We additionally are focusing on determining the concentration of surface ions using conductometric titration, whereby conductivity measurements are taken during titration.

G. Tucker, B. Brasseur, P.V. Schwartz; 'Surface tethered DNA provides a template for subsequent positioning of nanostructures', Presentation at COSAM, 2006.

We experiment with using the sequence selectivity of DNA-hybridization to guide the assembly of microspheres onto a templated surface. The surfaces can be templated by means of rubber stamping of Mercaptohexadecanoic Acid (MHA), which terminates in a COOH group which subsequently reacts with a short strand of DNA. We also explore more precise patterning of MHA by means of lithography with an Atomic Force Microscope (AFM). The DNA templated surfaces can then selectively bond microspheres

## **Large-scale distributed wireless networks**

Project Investigators:

John Seng & Diana Franklin  
Computer Science  
California Polytechnic State University  
San Luis Obispo, California

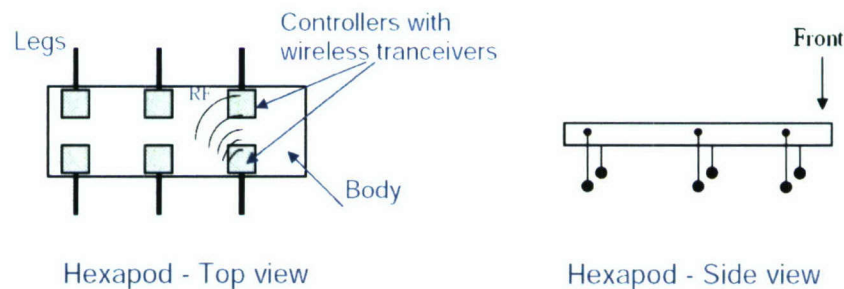
## Abstract

This project investigated the process of developing an experimental robot that utilizes a wireless network of microcontrollers for distributed motion control. The study successfully built and tested a proof-of-concept legged robot using a distributed wireless control strategy, and implemented directional walking algorithms for 4 and 6 legged robot designs. Such a robot design may prove applicable to rough-terrain wheeled and legged robots. Through this effort, this project provides an initial test platform and lays groundwork for future study of this control approach.

## Introduction

Our work in the area of wireless sensor networks has been in two areas: first, to develop protocols for a short-range wireless sensor network to control a legged robot and second, to investigate using wireless sensor network nodes coupled with accelerometers. We will report on the short-range wireless sensor network robot design.

The concept behind this study was to investigate the use of a short-range wireless sensor network as a controller for a legged robot. Such a robot would have a wireless sensor that would control each leg, and the leg controllers would coordinate motion via wireless communication. A robot design where each leg is controlled by a separate controller can be very useful in applications where: the robot needs to tolerate leg failure, the design requires that the legs be physically separated by a large distance where wires are not practical, and the robot design requires some modularity in which legs need to be added or removed. A conceptual drawing of such a design is shown in Illustration 1. Each leg is controller by a microcontroller and all of the microcontrollers are able to communicate via wireless.



*Illustration 1: Block diagram of robot*

In this project, we have achieved the following:

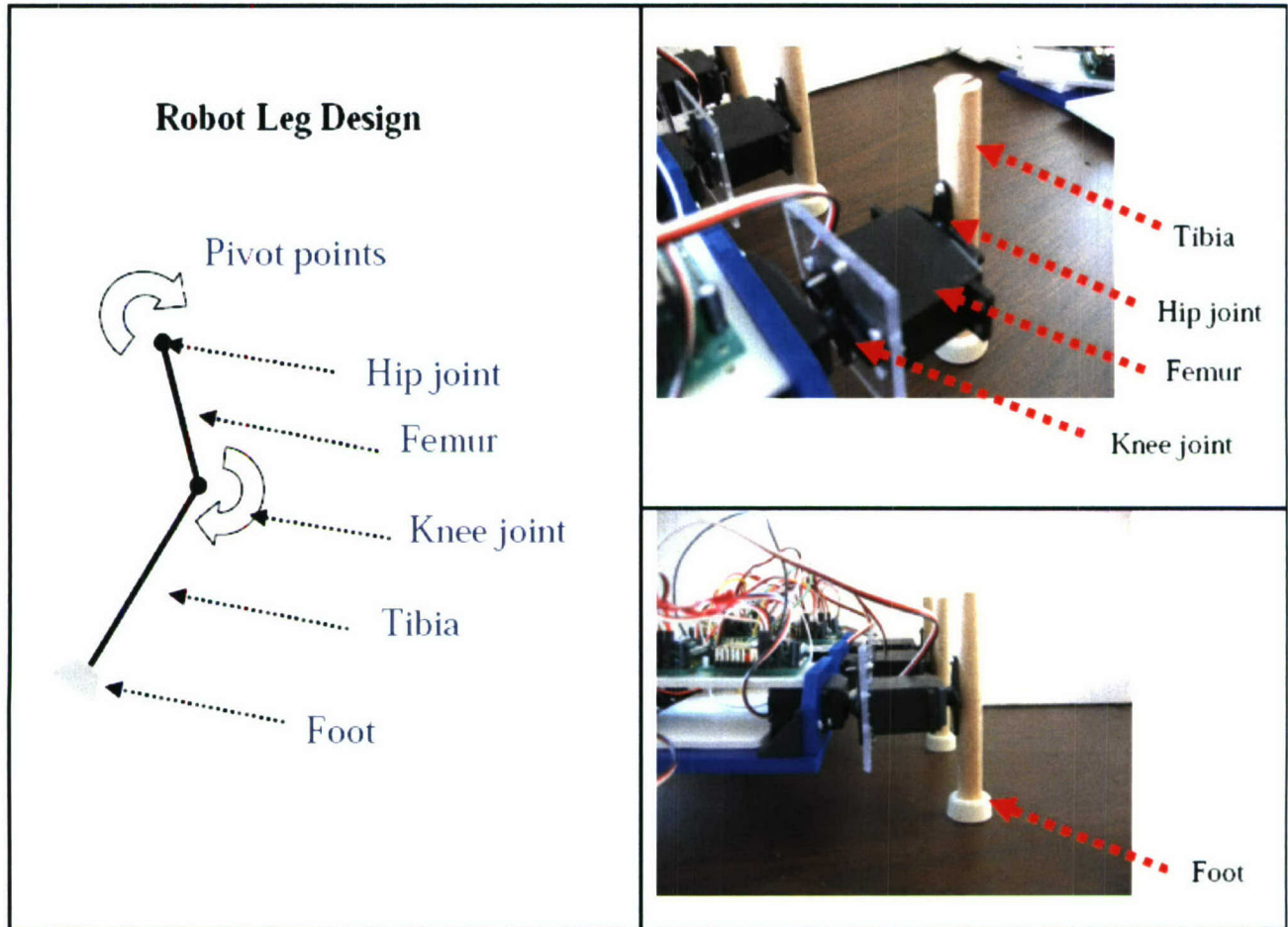
- Build a physical 4-legged and 6-legged robot with one controller dedicated to each robot leg. The robot is to serve as a proof of concept and test platform for future study.
- Design and implement walking algorithms for 4-legged and 6-legged gaits.
- Investigate wireless communication schemes to generate 4-legged and 6-legged walking patterns.
- Investigate fault tolerance and recovery approaches for a six legged robot in the presence of leg failures.

## Hardware Design

We have implemented the wireless controller modules that control each of the legs. The wireless controller modules operate in the 2.4GHz frequency range and provide high-bandwidth, low-latency network connectivity. The wireless modules are controlled using an Atmel ATMega32 running at 8MHz. Using the wireless controller modules, we have developed protocols which allow a master controller node to control the position each of 2 servo motors for each leg. This is done by having the master node send commands to each of the other slave nodes. The data for the servo positions is packetized in a protocol format we defined.

We have implemented 4-legged and 6-legged designs. The concept behind our design is to have a robot where multiple legs can be added or removed. The wireless controllers allow for this capability. At this point in time, we have tested 4-legged and 6-legged designs. Illustration 2 shows the basic design of each of the legs.

In addition, we have looked at fault tolerant approaches in the presence of leg failures. Should a leg design fail in the robot, it is desirable that the robot gracefully degrade. The controllers should be able to detect the failure of one or more legs because of various reasons (running out of batteries, failed motors, etc.) and should be able to notify a master controller node. The master controller nodes should adjust the gait accordingly. We have implemented some fault tolerance into the 6-legged robot. It is capable of detecting failures in 2 of the 6 legs. When such an event occurs, the robot degrades into a mode where it continues walking on 4 legs in a gait similar to the 4-legged robot design.



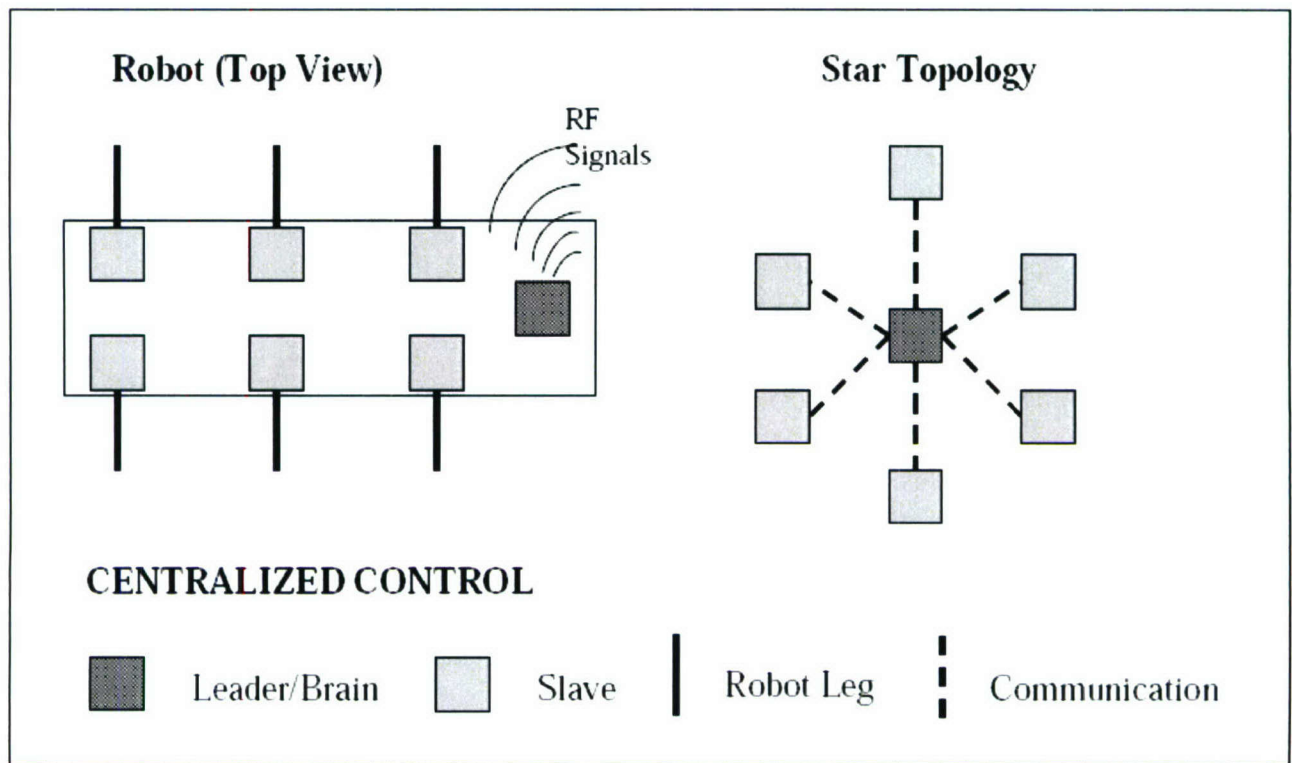
*Illustration 2: Leg design*

### **Software Design and Communication Models**

To walk forward, the leg controllers must communicate to coordinate leg movements. Since each leg controller is equipped with a wireless transceiver, each can be viewed as node in a wireless network.

This research work considered two wireless communication models to generate the 4-legged and 6-legged gaits—centralized and decentralized. In the centralized model as shown in Illustration 3, a single node acts as a leader node and sends move commands to each robot leg. The leader node can be on-board the robot or can be located off of the robot and need not be attached to a leg; it effectively serves as the coordinator of the robot leg motion. In the decentralized model, any node can assume leadership and direct the movement of the other legs.

Communication between nodes may be one or two-way. In one-way communication, the leader node sends a move command and does not wait for an acknowledgment (ACK) from the slave node. In a two-way communication, the leader node waits for an acknowledgment from the slave node before it communicates with another node in the network.



*Illustration 3: Diagram of control model*

The centralized model is analogous to a star network topology. The leader node broadcasts servo coordinates to all other nodes in the network, while the other nodes passively listen and optionally return an acknowledgment. Communication is limited to messages that travel between the slaves and the leader node; slave nodes do not exchange messages between each other. Message delivery and decision making are therefore the responsibility of the leader node.

There are advantages and disadvantages to this model. The advantages are:

- Software on the leg controller board is simple. It merely waits for servo positions from the leader-brain and moves the leg according to the received coordinates.
- Because control is centralized to the leader-brain, if one leg fails, the others are not affected so there is potential for recovery. On the hexapod, if one or both middle legs fail, the robot can assume the quadruped diagonal gait and walk on 4 legs.

The disadvantages are:

- An obvious downside to this model is that if the brain or leader fails, the entire network is destroyed and the robot is disabled with no potential for recovery.

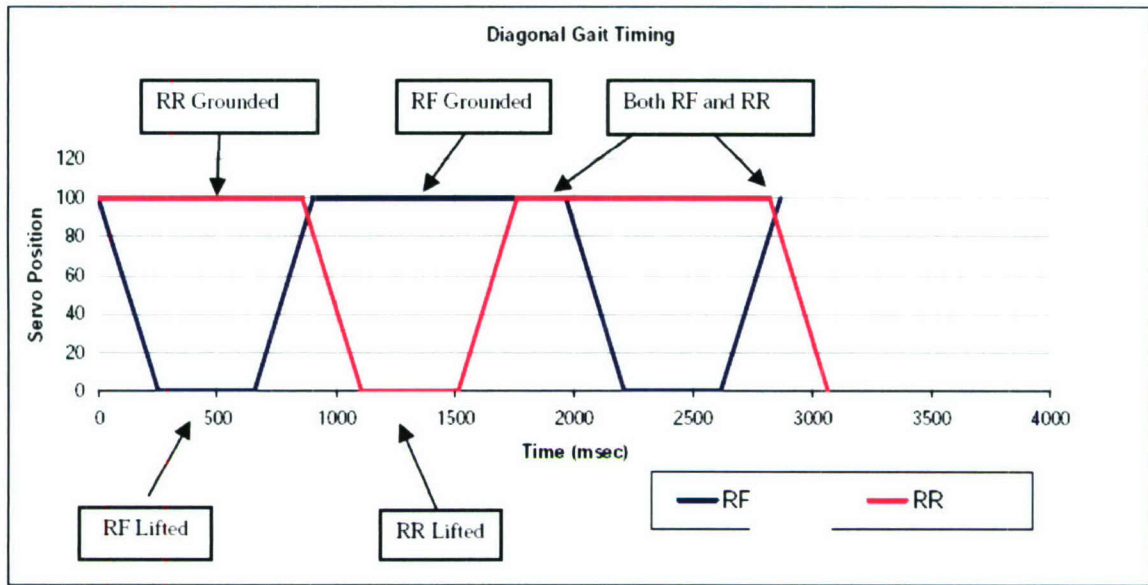
In this work, we focus on the centralized model. This centralized model was successfully implemented on the robot to produce both 6-legged tripod gait and the 4-legged diagonal gait.

### Leg Motion Timing

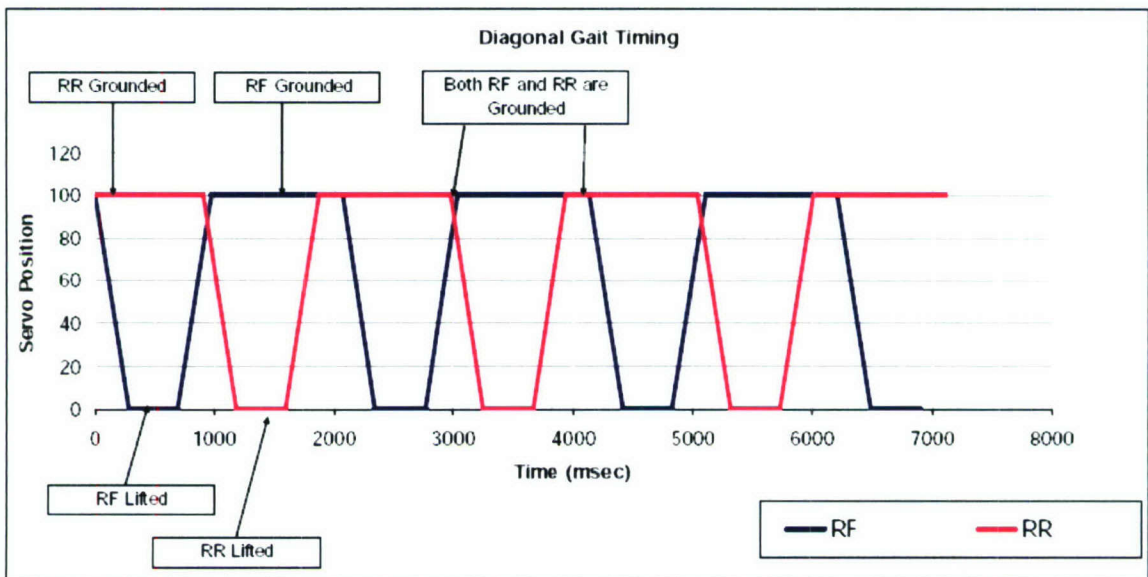
The figures in Illustration 4 shows the timing necessary for our functioning robot gait for the 4-legged robot design. Abbreviated notation is used to refer to each of the 4 legs: RR refers to the right-rear leg; RF refers to the right-front leg. Not shown is the timing for the left-rear and left-front legs as it is quite similar to that of the right-rear and right-front legs. The Y-axis refers to the servo position which is an 8-bit value corresponding to the position where the leg joint motor is currently located. The X-axis refers to the elapsed time. The figure shows the timing required to implement a successful 4-legged diagonal gait for our robotic design.

Figure (a) shows the timing required for proper motion of the robot when there is no packet loss (i.e. all packets sent from a microcontroller are successfully received by the destination microcontroller).

Figure (b) show the timing required when there is packet loss. In this case, it is assumed that all packets are lost initially and a retransmit of the packet is required. In this case, each communication transaction incurs a timeout delay of the first packet lost. We found that the delay was relatively insignificant and that our short range wireless network worked well. In practice, we found that the packet drop rate was ~3%. Under these conditions, using a short range wireless network provided near-instant reaction times for our robot and makes the transition from a wired robot to a wireless design very usable.



(a) Ideal Case: No Packet Resend



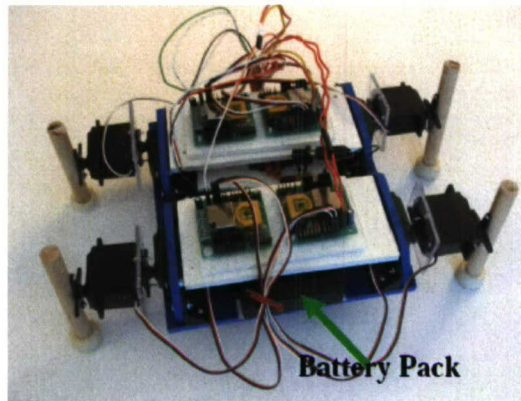
(b) Worse Case: Maximum Packet Resend

Illustration 4: Leg Timing

### Final Implementation

Illustration 5 shows a picture of our final design. This is an implementation of the 4-legged robot. In this implementation, the legs are powered from the same battery pack for testing purposes. In an full implementation, the legs and microcontrollers could easily be power from

separate battery packs which would realize fully independent legs in terms of power and control.



(c) Revision 3 - Final Design

### Summary and Conclusion

This work set out to develop a robot with a control system consisting of several distributed wireless microcontrollers. This goal was accomplished with the completion of the following tasks:

- Built a physical experimental robot system with distributed legs as a proof of concept and test platform. The robot walks according to pre-programmed/pre-recorded movements and can be configured into a quadruped or hexapod.
- Designed and implemented unidirectional walking algorithms for the 4-legged and 6-legged experimental robot. Walking is successful on a flat surface.
- Investigated fault tolerance and recovery approach for the six-legged robot in the presence of leg failures. In the recovery implemented, the hexapod can assume the 4-legged diagonal gait if one or both of the middle legs fail.
- Implemented the centralized communication model to generate the quadruped diagonal and hexapod tripod gaits.

**Passive directionally sensitive laser Doppler velocimeter**

Project Investigator:

John Sharpe  
Physics Department  
California Polytechnic State University  
San Luis Obispo, California

The purpose of this project was to implement a passive, directionally sensitive laser Doppler velocimeter (LDV). The achievements of the project include:

- The proposed concept has been proved and we have shown that a directionally sensitive laser velocimeter can be implemented using wavelength/polarization multiplexing. However, a problem we encountered was the difficulty dealing with dispersion and, especially, the wavelength shifting when using diode lasers.
- We have researched miniaturizing the system and have entered into a non-disclosure agreement with a company (Measurement Science Enterprise (MSE) [www.measurementsci.com](http://www.measurementsci.com)) that manufactures miniature LDV systems.
- After consultation with MSE, including a follow-up visit to their manufacturing facility in Pasadena, we think that the market for our concept might be too small for continued development for miniature systems at this time. In part this is due to relative complexity of our system making it harder to manufacture. MSE has also pointed out that in passive LDV systems like ours there are problems in obtaining the necessary dynamic range in the processing electronics.
- In conjunction with MSE we have developed several new concepts in relation to optical frequency shifting based on modern piezo and liquid crystal technology.

This work was included in the senior projects of the following students:

Student name	Project Title	Current status of student
Adelsheim, Michael (Phys)	Direction sensitive laser velocimeter	Accepted, University of California, Hastings Law School.
Elton, Mark (Phys)	Signal acquisition and processing for a laser velocimeter	Accepted, PhD Program, Mechanical Engineering, Georgia Tech.

Our project combined the recent proposal of Büttner and Czarske [Buttner2003] with a well-known technique of polarization shifting to essentially create two optical fringe patterns which are projected into the flow. The two patterns have the same period, but are generated from different wavelength sources and are held apart by a quarter of a period. Depending on the direction of the flow, the signal at one of the wavelengths either leads or lags the other. Figure 1 shows a schematic of the optical system while figures 2 and 3 are photographs of the overall optical setup and details of the detection system, respectively. Figure 4 shows typical data taken from the system when the “flow” (in practice we use a ground glass scattering plate on a motorized translation stage) moves in the forward and reverse directions. The change in phase lag as the direction changes can be clearly seen.

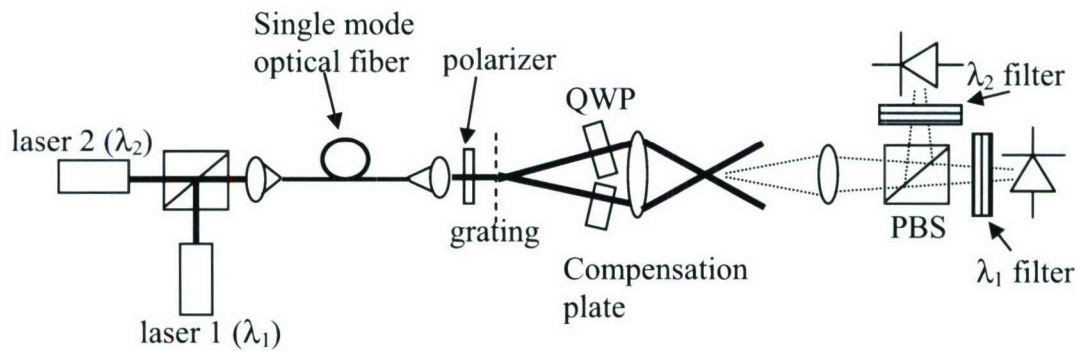


Figure 1. Schematic of the setup to test passive directional discrimination using polarization and wavelength encoding. The light from the two lasers is combined at the input beamsplitter, collimated and directed to the grating. The  $\pm 1$  orders are selected and one order passed through the quarter-wave plate, QWP. The other passes through a compensation plate to equalize the paths and ensure high contrast fringes when using a low-coherence laser diode. It was found that slight tilts of the compensation plate could be used to bring the two fringe patterns into precise quadrature alignment. After scattering, the light is collected and the orthogonal polarizations selected using the polarizing beamsplitter, PBS. Final selection of the wavelengths is done using the narrowband interference filters in front of the detectors.

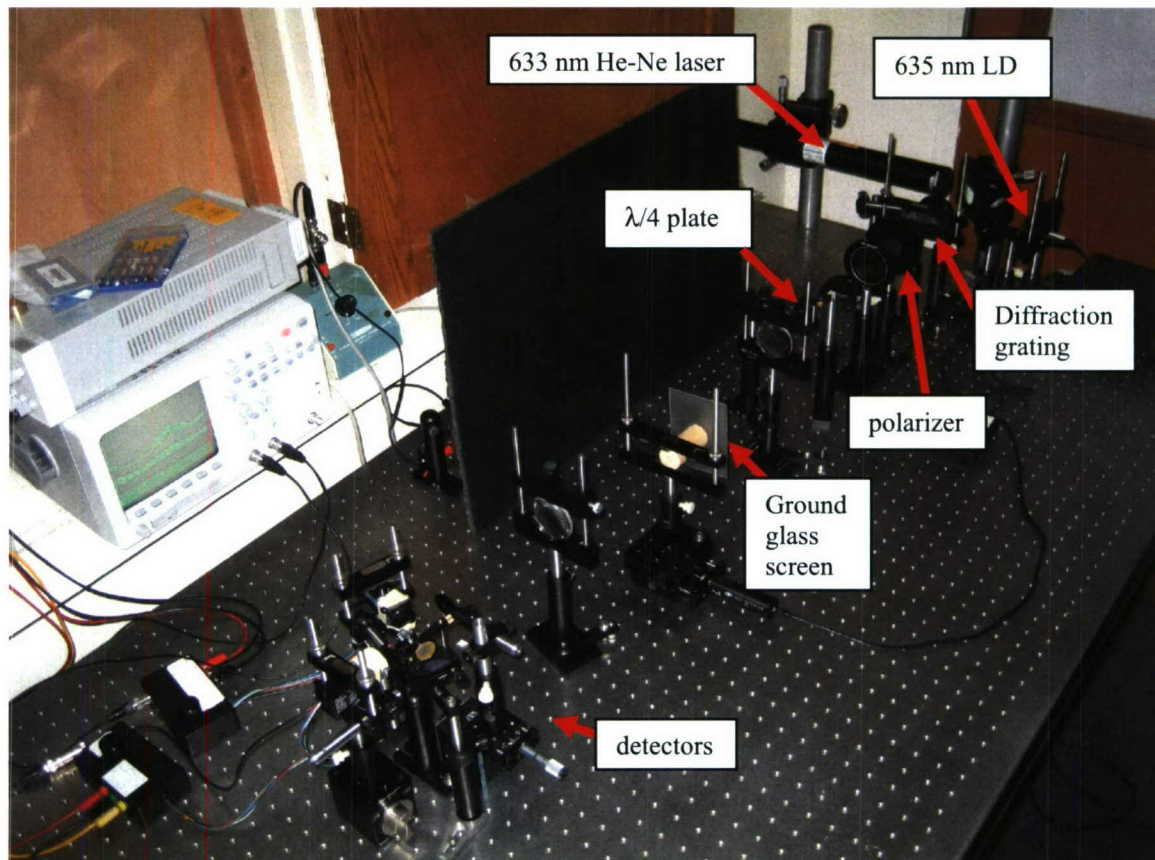


Figure 2: Laboratory setup

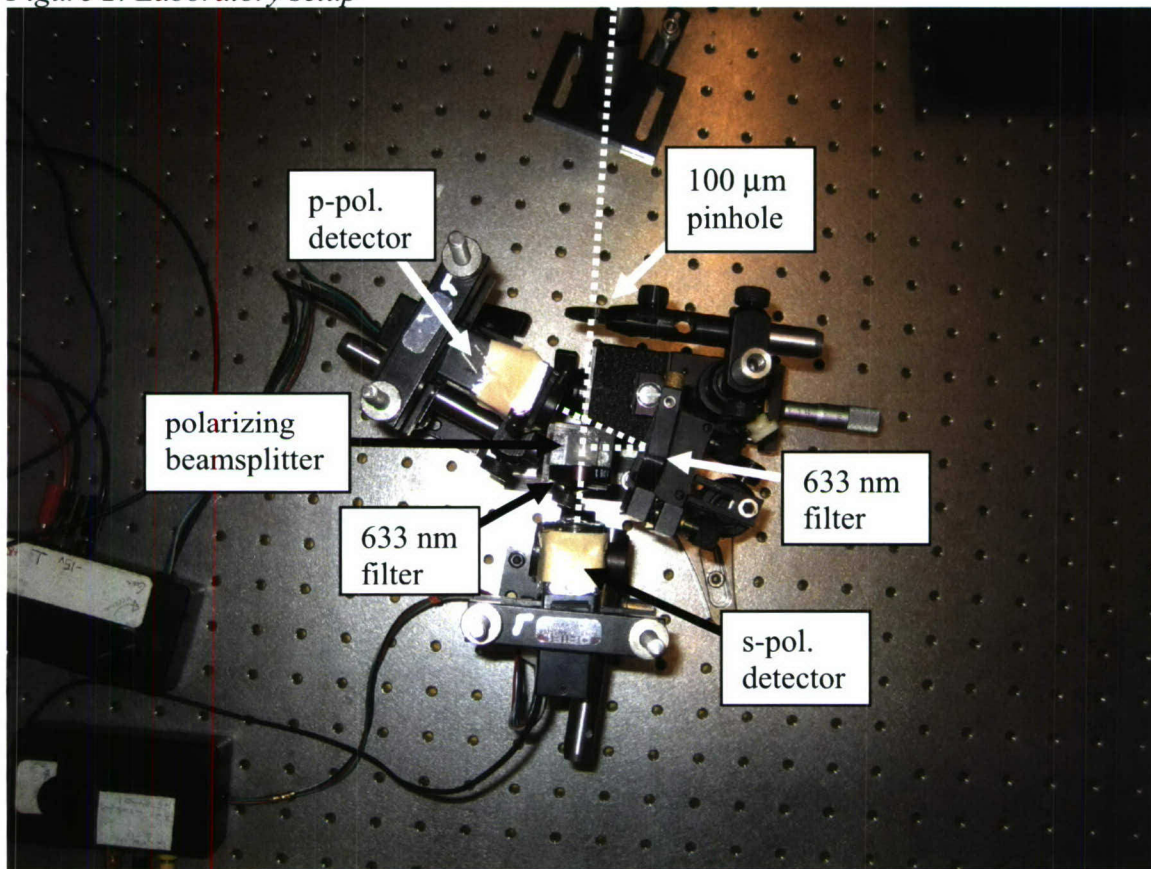


Figure 3: Detection system. The light from the scattering glass plate is imaged through the 100  $\mu\text{m}$  pinhole and is divided into horizontal and vertical polarization components using the polarizing beam splitter. The horizontally polarized component at 633 nm is transmitted by the 633 nm filter to the s-polarization detector. The vertically polarized component at 635 nm reflects from the other 633 nm filter and goes to the p-polarization detector.

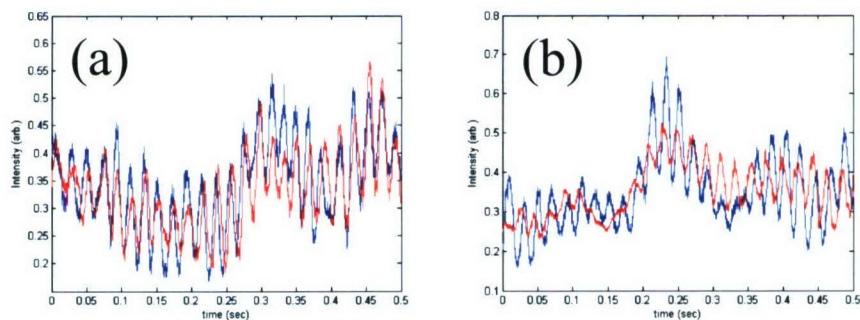


Figure 4: Photodetector signals for 633 nm (blue) and 635 nm (red) wavelengths. (a) Signals obtained with the scattering screen moving in the “forward” direction. (b) Signals obtained with the scattering screen moving in the “reverse” direction.

The phase shift between the two signals was determined by cross-correlation and we found that over multiple runs the phase differences were always within 0.2 radians of the expected  $\pm \pi/2$  phase lag. Figure 5 shows results of a series of runs with the ground glass screen moving in the “forward” and “reverse” directions.

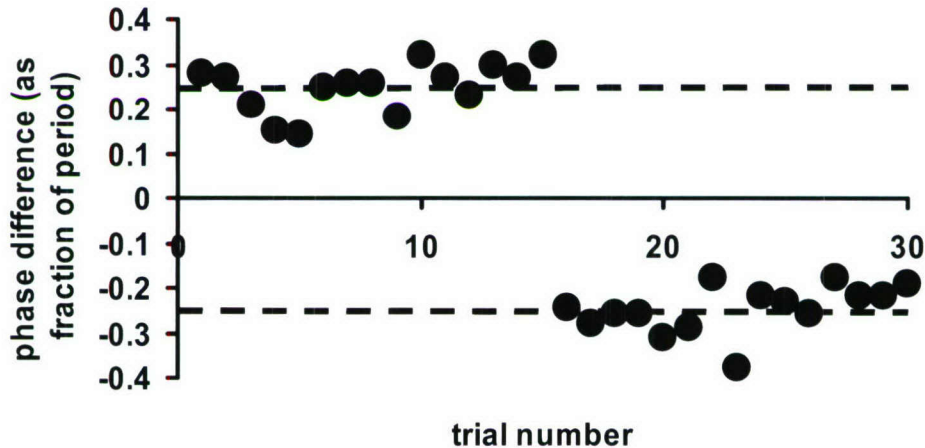


Figure 5. Phase difference between the signals from the two lasers for thirty different trials. In the first fifteen trials the screen moved in the forward direction. In the rest of the trials the screen moved in the reverse direction. The dotted lines mark the expected phase shift of  $\pm$  quarter of a period.

The data of figure 3 was taken over a particularly stable time of the diode laser. We found that even though the laser was placed in a thermoelectrically cooled holder which certainly mitigates the wavelength drift, the hops of the laser diode were enough to switch the phase between the two fringe patterns to a random value. This problem could indeed be mitigated by either using another (non-diode) laser or stabilizing the laser diode. However, both of these options would greatly increase the cost of the setup.

After we had communicated our findings to our industry collaborator, we were also made aware of a technical point that had not been discussed in the literature. This is that in our passive system all of the generated signals are centered at zero hertz. In contrast, the active shifting schemes are centered at some higher frequency (typically from 0.1 – 100 MHz). It turns out that it is easier to design the light detection and processing electronics to work around a non-zero center.

We considered the packaging of the system into a miniature probe head and this is schematized in figure 4. The components required for this system are either readily available (e.g. ball lenses, miniature prisms) or could be supplied at relatively low cost from current manufacturers.

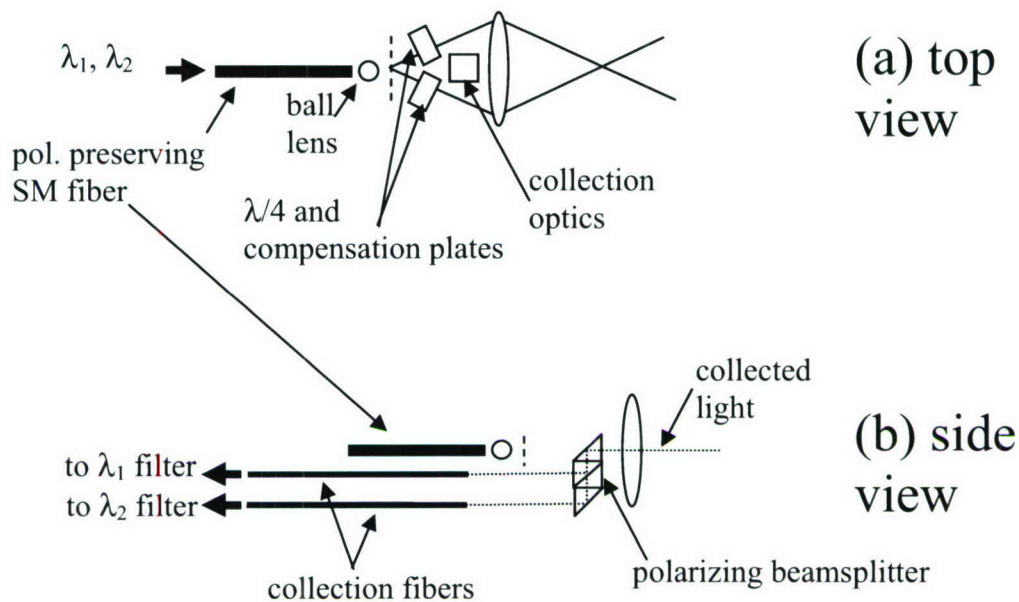


Figure 4: Schematic for an implementation of a velocimeter probe head. The top view, (a), shows the light delivered to the probe head already polarized. This light is collimated using the ball lens and divided using the grating. After passing through the quarter wave and compensation plates the beams are combined at the measurement volume. The scattered light is collected using the delivery lens and fed to the collection system, shown in the side view (b). The orthogonal polarization components are separated using the polarizing beam splitter and delivered to the wavelength filters using the collection fibers.

Again, following discussions with MSE it was decided that it would probably not be cost-effective to pursue this approach to LDV at this time. There are already probes available which, although they have design problems associated with the frequency shifting component, are fairly reliable and are quite cheap to manufacture.

#### Reference

Buttner, L. and J. Czarske. "Passive directional discrimination in laser-Doppler anemometry by the two-wavelength quadrature homodyne technique." *Applied Optics* 42 (19) 3843-3852 (2003)

**Nano-grain cutting tool for machining high-temperature  
aerospace materials**

Project Investigator:

Daniel Waldorf  
Industrial & Manufacturing Engineering  
California Polytechnic State University  
San Luis Obispo, California

## Summary

As the aerospace industry attempts to save money by implementing high-speed machining of parts and components, the costs and limitations of commercially available cutting tools represent a serious challenge. Furthermore, the new materials being used, including titanium alloys and superalloys are even harder and more abrasive than traditional aircraft aluminums. The materials demand slower cutting speeds and result in shorter tool lives unless expensive (and brittle) ceramic tools are used. A new cutting tool formulation and fabrication method has been developed and patented with the potential to greatly increase the utilization of less expensive “carbide” tooling to machine these materials. The new composite tool material uses tungsten carbide (as do most carbide tools) and several added ceramics as the reinforcing materials but improves the overall properties by introducing rhenium and nickel-based superalloy within the metal binder. The new binder materials are possible because of a newly developed powder sintering process for economically producing the tools. The resulting product is formed with a nano-grained material structure to achieve a harder, tougher base material that retains its hardness at elevated temperatures. In addition to preliminary material tests that show distinct promise and advantage, controlled machining tests were conducted to investigate the effect of binder composition on cutting tool life. The tests showed a clear advantage of the new formulations. Tools with rhenium and superalloy as the binder metals showed a tool life roughly twice that of the tools with only cobalt in the binder. A cost analysis of the new tools indicates that they would cost roughly 10-50% more than cobalt-only tools to produce.

## Plan of Work

The plan of work involved two key efforts:

- Produce a set of tooling samples with varying metal binder content, and
- Conduct a series of experiments to test the wear-resistance of the tools under actual conditions of machining

24 different grades of tooling were designed and 24 lots (650g each) of powder were made. The 24 lots are list in Table 1. R-95 and U-720 are two different Ni-Based Superalloys. T3.5 and T32 are WC powder with the average grain size of 0.55 and 2  $\mu\text{m}$ , respectively.

Production of these samples involved a multi-step processing sequence. Raw powders of tungsten carbide and other carbides (TiC, TaC, VC), cobalt, rhenium, and nickel-based superalloy were purchased from an outside vendor. Milling (to reduce grain size), drying, and sifting the powders took place to create the right blend for each of the 24 lots of powder. The powder was then cold-pressed into bars under high pressure to form the initial “green” compacts. The bars were then presintered in a conventional oven just hot

	Re	R-95	Co	U-720	T3.5	T32	TiC	TaC	Cr <sub>2</sub> C <sub>3</sub>	VC		Density
	V1				V5		V7			V10		
	%	V2%	V3%	V4%	%	V6%	%	V8%	V9%	%	vb%	g/cc
d1A			10		89.5	0			0.3	0.2	10	14.923
d2A			10			90					10	14.97
d3A			5		94.5	0			0.3	0.2	5	15.253
d4	9	0.5		0.5	0	89			0.7	0.3	10	15.945
d5	9	0.5		0.5		89					10	15.945
d6	4.5	0.25		0.25	0	94			0.7	0.3	5	15.741
d7	5	2.5		2.5	89	0			0.7	0.3	10	15.431
d8	5	2.5		2.5		90					10	15.525
d9	2.5	1.25		1.25	94	0			0.7	0.3	5	15.484
d10	1.5	4.25		4.25	0	89			0.7	0.3	10	14.982
d11	1.5	4.25		4.25		90					10	15.075
	0.7			2.12								
d12	5	2.125		5	0	94			0.7	0.3	5	15.259
d13	1.5	3	2.5	3	89	0			0.7	0.3	10	15.004
d14	1.5	1.75	5	1.75	0	90					10	15.119
	0.7											
d15	5	1.5	1.25	1.5	94	0			0.7	0.3	5	15.27
d16	9	0.5		0.5	0	83	3	3	0.7	0.3	10	15.59
d17	4.5	0.25		0.25	0	88	3	3	0.7	0.3	5	15.386
d18	5	2.5		2.5	83	0	3	3	0.7	0.3	10	15.076
d19	2.5	1.25		1.25	88	0	3	3	0.7	0.3	5	15.129
d20	1.5	4.25		4.25	0	83	3	3	0.7	0.3	10	14.626
	0.7			2.12								
d21	5	2.125		5	88	0	3	3	0.7	0.3	5	14.904
d22	1.5	3	2.5	3	83	0	3	3	0.7	0.3	10	14.648
d23	1.5	1.75	5	1.75	83	0	3	3	0.7	0.3	10	14.67
	0.7											
d24	5	1.5	1.25	1.5	88	0	3	3	0.7	0.3	5	14.915

Table1. Composition of 24 powder lots

enough to start the particle bonding that allows significant handling of the compacts without risk of damage. 4 test bars for each lot were pressed and sintered in this manner. 24 lots were approved.

The samples were then pressed into the shape of cutting tool inserts (industry designation SNU433), with 30 inserts for each lot. Next, the newly developed solid-phase sintering process was performed at Genius Metal. This sintering is also performed in a conventional oven but does not reach the melting temperature of the metal matrix material (unlike with traditional sintering of cobalt-based carbides). The purpose is to further densify the samples and begin the real bonding between all of the constituent

particles. After this first sintering stage, the samples were submitted to a hot, isostatic press (HIP) to complete the bonding and densification. HIP involves placing the samples in a graphite die (must be able to withstand very high temperatures) and submerging in a high-temperature, high-pressure gas (typically Argon) medium. Here, the further physical compaction and the high-temperature sintering happen simultaneously to achieve the final material composition. 22 lots were completed with this process and were sent on to the next stage. The remaining 2 lots (d4, d5) experienced problems with pressing and viable samples were not obtained.

Lastly, the samples were sent to a machine shop to grind the final surface, profile, and edge condition. Surface roughness is critical to fracture toughness in cemented carbides. Similarly, a smooth profile (e.g., corner rounding on square inserts), and appropriate edge condition ensure that premature fracture is avoided. For the hard, potentially brittle samples developed here, a rounded hone combined with a chamfered “T-land” was put on each insert to provide the best defense against chipping and fracture. Additional edge hones without the T-land were ground onto several of the insert samples.

While the tooling samples were being prepared at Genius Metal, several preparations for the machining experiments took place at Cal Poly. First, a capable student assistant was recruited and hired. The student has considerable experience in basic machining, CNC machining, and analysis of cutting tools, especially the hard ceramic coatings on tools. In addition, three Cal Poly engineering students were hired and trained on the procedures for running the experiments.

The students, together with the PI, developed a procedure for workpiece preparation (i.e., surface cleaning, machining of support mounts), machine set-up (workholding and tooling selection), machine programming (CNC code), test conditions (two cutting speeds with set depth of cut and feed rate), data collection, and inspection. For data collection, a form was developed to easily document the cutting conditions and results from each test sequence. A test sequence consists of four to five cutting tests on the same tool sample (note that each sample contains up to eight possible edges for testing). The first test will run an edge to failure (based on wear scar, surface degradation, or fracture), with the remaining tests run to just 80%, 60%, 40%, and 20%, of the previously determined tool life. That way, a clear trend is recorded to determine the wear rate, rather than just the tool life. For inspection, the students and the PI collaborate to develop a program on the newly acquired OGP Smartscope 250 automated vision inspection system. The system uses machine vision (camera) and a laser to record four or more separate measures of the tool wear. Different cutting conditions and tool samples will typically cause different modes of tool wear, hence the different measures. The laser measurement of the depth of the wear crater on the rake face is state-of-the-art and will be a valuable tool for reporting and publishing results. Machining tests have just begun as of this progress report.

The students and PI also developed a procedure for taking cutting force measurements of a select subset of the machining trials. The cutting force measurement requires a separate tool dynamometer mounted on a different machine than the standard trials. However, force data can give additional insight into the mode and rate of wear occurring on a

particular tool and can provide a further basis of comparison amongst tool types. When needed, force data is recorded on the same data collection form as noted above.

## Results

**Tool Life.** The turning test results on Inconel 718 are summarized in Figures 2 and 3 below for cutting at speeds of 50 and 75 mpm. They show that tool life (judged based on a flank wear land width of 0.5 mm) is dramatically affected by the binder composition. Of the three tool grades that contained only cobalt in the binder (essentially the “control” condition since it matches a standard carbide tool), all performed essentially the same and are well-represented by tool d3A. All of the other tool grades (having alternative binder materials) performed better than the three cobalt-only grades. Several typical grades were selected and placed on the graphs. As can be seen in the graphs, tool life at V=50

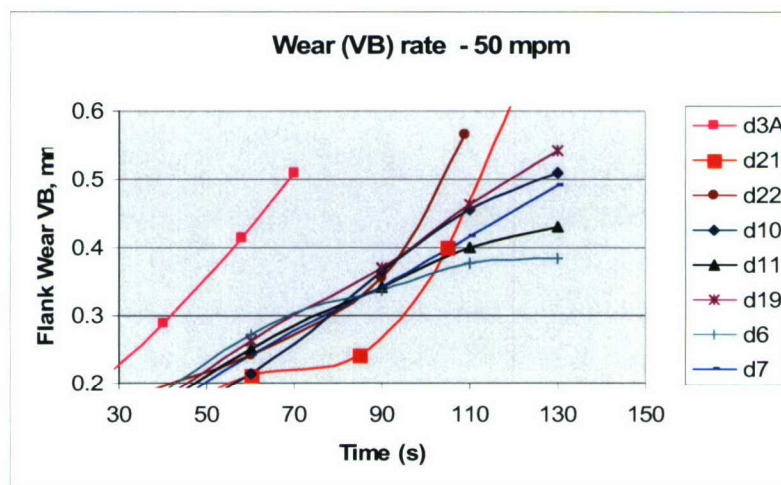


Figure 2 Flank Wear vs Time for Cobalt-Only (d3A) and Alternative-Binder Tools (DOC= 2 mm, Feed=0.1mmpr) at V=50 mpm

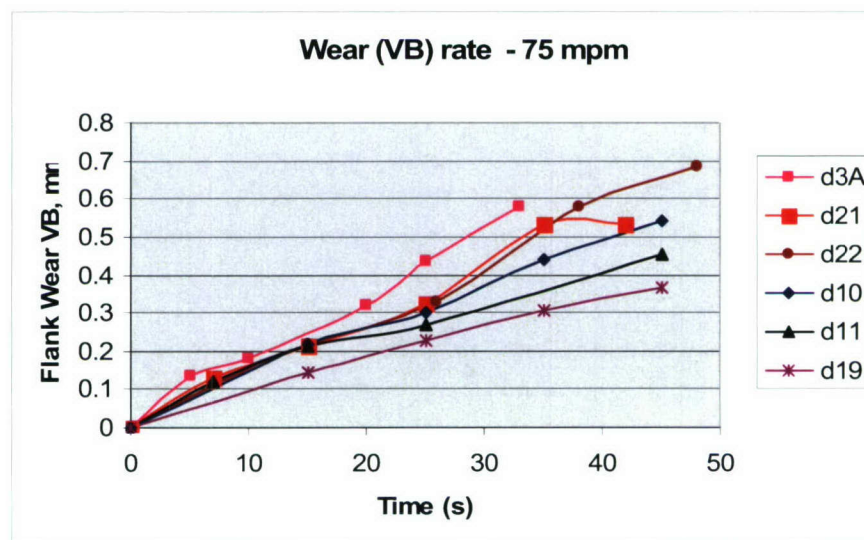


Figure 3 Flank Wear vs Time for Cobalt-Only (d3A) and Alternative-Binder Tools (DOC= 2 mm, Feed=0.1mmpr) at V=75 mpm

Alternative-Binder Tools (DOC= 2 mm, Feed=0.1mmpr) at V=75 mpm mpm is approximately doubled (for VB=0.5 mm wear) for the alternative-binder tools; i.e., 100% increase compared to the d3A (cobalt only). At higher speed (V=75 mpm), the tool life is increased 20 to 100%. It should be noted that some the tests (i.e., for tools d10, d11, d19) are based on a single set of results and could not be replicated due to limited tool samples.

The only evaluation of crater wear was completed for the three cobalt-only inserts, 2 at V=50 mpm, and one at V=75 mpm. These are for non-chamfered inserts since the crater wear is much easier to measure for these than on the chamfered insert edges. Results are shown in Figure 4 below. Note that crater wear was not the dominating wear mechanism for any of these tests.

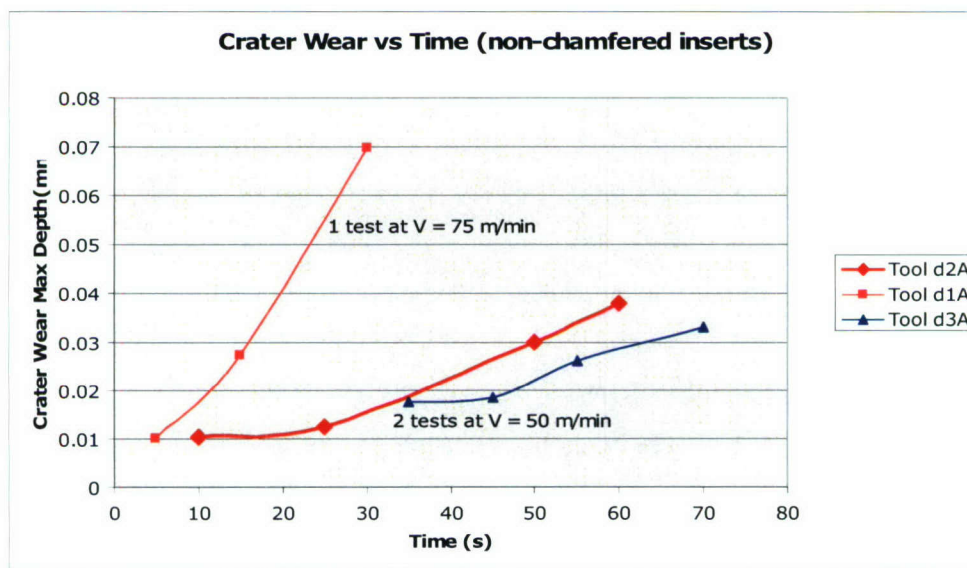


Figure 4 Crater wear depth vs Cutting Time for Various cobalt-only binder grades

A few quick comparisons were also made against Greenleaf WG-300 SiC whisker-reinforced alumina inserts. Two comparisons were made: one at the recommended conditions for ceramics (i.e., V=300 mpm, depth=0.23 mm) and one at conditions similar to the above tests (i.e., V=75 mpm, depth=2.0 mm). Table 2 shows the results of the high-speed tests. The ceramic insert wears normally, while the d22 insert (and other carbides tested) fails rapidly from excessive flank wear/deterioration.

Time of Cut (sec.)	Flank Wear Ceramics (mm.)	Flank Wear d22 (mm.)
45	0.05844962	Failed
90	0.109982	Failed

Table 2 High Speed Comparison of WG-300 vs d22 Carbide

Table 3 shows the results at the lower speed. In this case, fairly comparable flank wear is seen for all of the inserts. All feeds were 0.1 mmpr.

Time of Cut (sec.)	Flank Wear Ceramics (mm.)	Flank Wear d22 (mm.)	Flank Wear d21 (mm.)
7	0.0934974	0.0832358	0.0912368
25	0.1839722	0.2212213	0.2107184

Table 3 Lower Speed Comparison of WG-300 vs d21 and d22 Carbide

Cutting Forces. The students and PI measured cutting forces during a select subset of the machining trials. The cutting force measurement required a separate tool dynamometer mounted on a different machine than the standard trials but can give additional insight into the mode and rate of wear occurring on a particular tool. Figures 5 and 6 show the two main force components (Cutting and Thrust) plotted against wear width for several of the tool grades produced and for one commercially purchased Valenite VC-2 grade. The graphs show that the forces increase about the same amount, regardless of which tool is being used. The cutting force increases about 30% and the thrust force increases about 250% as the wear land grows to .4 mm in width. In order to investigate the effect of the edge condition, a set of tests was also run comparing cutting forces for tools that were not chamfered during the final geometry grinding stage as described above. The un-chamfered tools resulted in approximately the same cutting forces as the chamfered tools.

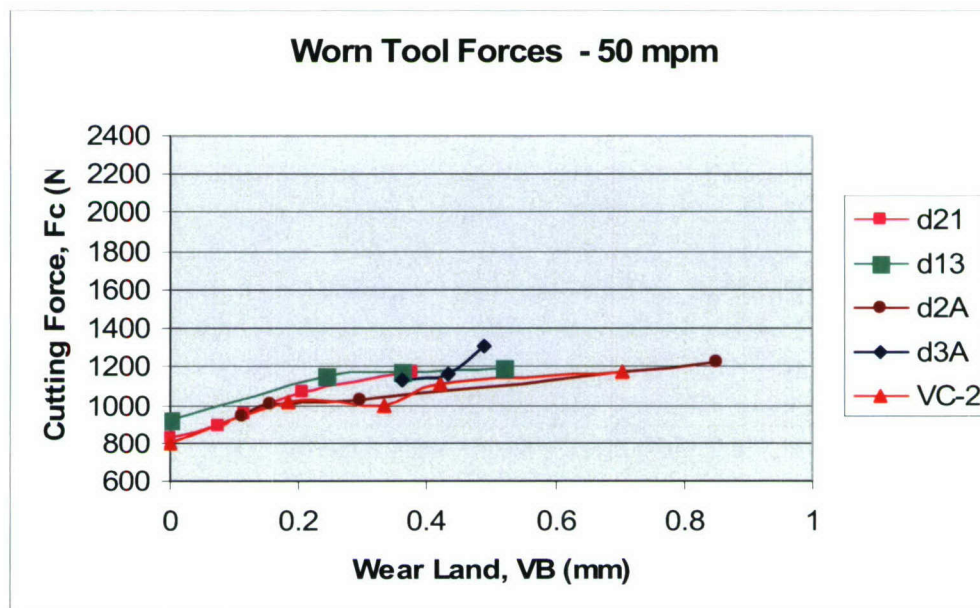


Figure 5 Cutting Forces vs Wear Land (DOC = 2 mm, Feed = .1 mmpr) for V=50 mpm

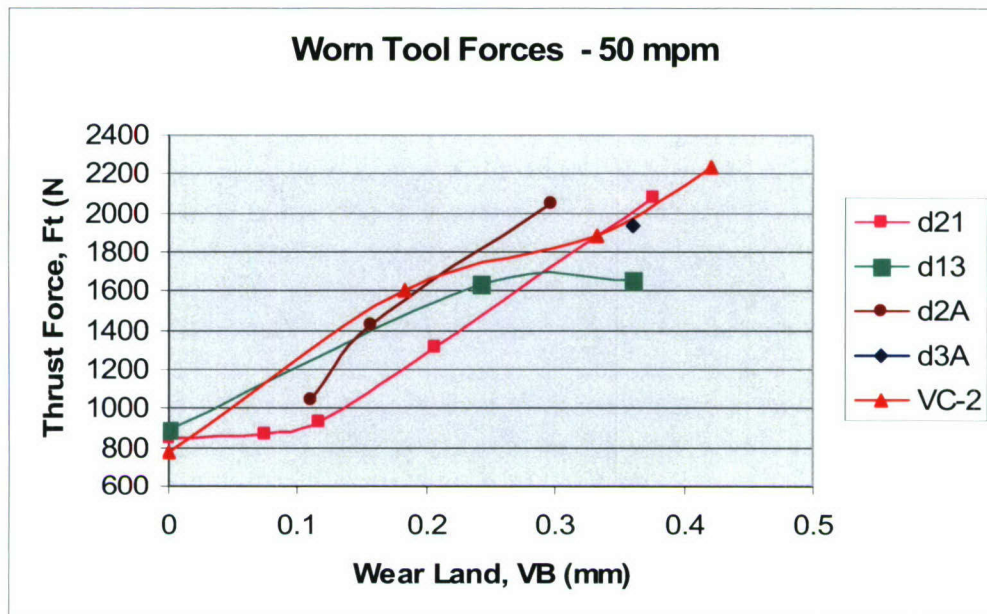


Figure 6 Thrust Forces vs Wear Land (DOC = 2 mm, Feed = .1 mmpr) for V=50 mpm

Material Constituents. Among the various material factors examined in the test matrix, there appeared to be a dominant effect only with binder % and with presence of TiC/TaC. Three sets of direct comparisons were made between tools with “Fine” grain size and those with “Course” grain size, but the results were evenly split as to which tools lasted longer. Of the nine direct comparisons made regarding “small” (5%) vs “large” (10%) binder percentage, none of the “large” samples outperformed the “small” samples at 50 mpm speed and only two lasted longer for the 75 mpm speed. In the eight direct comparisons between presence and absence of TiC/TaC, in only one case did the absence of TiC/TaC result in a longer tool life. Finally, five direct comparisons were made amongst samples with different percentages of Co, Re, and superalloy; unfortunately, the results showed no consistent trend in terms of superior performance. It is hoped that additional testing can be used to further optimize the relative proportions of these binder metals.

The comparisons across material constituents were made more difficult by several instances of inconsistent results. First, as the diameter of the Inconel workpiece grew smaller, its temperature increased dramatically since it had less ability to conduct away the heat from machining. Tool tests taken at this smaller diameter seemed to show much shorter lives and more irregular wear patterns on the tools. Second, also at small workpiece diameter, was the problem of chip control. At the small diameter, the chip was much more likely to get tangled in the workpiece causing chipping and additional wear on the tool and making measurement of wear more difficult. Third, it appeared from the tool samples themselves, that those produced early in the year (especially d12, d13, d14, d15, d20, d21, and d22) had less consistent structure (e.g., level of porosity or edge defects) than those produced later. Although attempts were made to randomize the

test sequence to minimize the effect of these factors, those efforts were not always successful.

**Cost Analysis.** An analysis of the production costs of the new tooling can be made based on recent purchases of raw materials, and a comparison can be made to currently available carbide and ceramic products. A standard 10-gram (SNG-432) indexable carbide cutting tool contains about 6% cobalt (i.e., .6 g) and 94% tungsten carbide (i.e., 9.4 g). Based on recent purchases of both powders, material cost would be \$.51 per insert (\$.075/g x 0.6 g Co plus \$.05/g x 9.4 g WC). Based on preliminary results above for d11 tool, a successful alternative-binder tool would have 93.3% WC (\$.47 for 9.33 g), 4.6% superalloy (\$.075/g x .46 g = \$.035) and 2.1% Re (\$5/g x .21 = \$1.04), for a total material cost of \$1.55 per insert, the extra Re essentially adding \$1 to the cost. Since standard grade WC-Co inserts typically retail for about \$8 per insert, the added Re in the d11 tool would increase the cost by about 12% (assuming other processing costs remain constant). For the other top performer, the successful d6 tool contains 94% WC (\$.47), .26% superalloy (\$.002), and 6% Re (\$2.98), for a total material cost of \$3.45 (a 40% increase in insert cost). Given that both of these tools showed more than two times the life of a cobalt-based tool, the modest cost increases would seem to be worthwhile.

Ceramic tools (e.g.,  $\text{Si}_w\text{-Al}_2\text{O}_3$ ) typically cost three to five times that of carbide tools; i.e., \$35 compared to \$8 per insert. Although the ceramics can cut at speeds of more than 300 mpm (compared to 25 - 50 mpm for uncoated carbides) on Inconel 718 and maintain a long life, they are recommended by the manufacturer only at light/finishing conditions (e.g., .1 mmpr or lower feed rate, .25 mm or lower depth of cut for the insert shape used in these tests – SNG-432). According to the manufacturer, anticipated tool life for the ceramic cutter at these conditions is about 5 minutes. By comparison, the carbides in this study ran at 50 mpm speed with .1 mmpr feed and 2.0 mm depth, which is comparable to the ceramic's material removal rate. The tool lives for the d11 and d6 tools were approximately 3 to 4 minutes, and this for *uncoated* tools cutting *without* coolant! If coatings and coolant can extend tool lives even modestly, these tools will perform at or better than the level of the ceramics (i.e., in terms of material removal rate) – at a fraction of the price.

## Conclusions

A set of new tungsten-carbide-based cutting tool samples has been produced using rhenium and nickel-based superalloy as the binder metal. These samples were used to cut Inconel 718 and their tool life was measured and compared to “control” samples that included just the standard cobalt as the binder. A brief comparison was also made with commercial ceramic inserts. Significant progress was made in the development of this new tooling material in addressing the machining challenges facing the aerospace industry.

Based on the testing completed for this project, the following conclusions can be drawn.

- The new tool grades that include rhenium or superalloy or both as binder metals performed significantly better than the cobalt-only binder tools. Although all of

these new tools showed increased life over the cobalt tools, several of the best performing grades showed approximately twice the tool life.

- Among the newly produced tooling samples, the presence of TiC/TaC and the use of a smaller (5%) binder percentage appeared to increase tool life by approximately 15% each. No clear results could be drawn about the optimal percentage of Co, Re, or superalloy or about the effect of grain size.
- All of the new tools run on Inconel failed by excessive wear on the tool flank, as opposed to crater wear or notch wear. This suggests that the increased hardness and hot hardness of the new tools was a predominant factor in explaining their longer life. Although crater wear was measured for several samples, the levels reached were not close to typical failure criteria for crater wear.
- A limited test of commercially available ceramic (SiCw-Al<sub>2</sub>O<sub>3</sub>) tools showed the advantages and disadvantages of the ceramics. At modest cutting speed (75 mpm, fairly high for carbides), the ceramic performed very similar to the newly created carbide tools; thus, given their high cost, ceramics performed poorly at this condition. At very high speeds (300 mpm), the ceramic held up and followed a pattern of normal tool wear while the newly created carbide tools failed immediately. However, the cutting load (i.e., feed rate and depth of cut) at the higher speed was drastically reduced, in order to protect the relatively brittle ceramic tools. Overall material removal rate was comparable between the high speed-low load tests and the modest speed-high load tests.
- Measurement of forces during cutting confirmed that the different tool grades (and a commercial carbide grade) had approximately the same forces applied to them during cutting under the same process conditions. It also showed that cutting with or without a chamfer resulted in no significant difference in cutting force.

Additional tests are ongoing with this project. The PI and research assistants have continued trying to replicate tests to find consistent results, especially using non-chamfered tools and cutting primarily on large diameter workpieces. Genius Metal will also recreate several of the tool sample grades in order to reproduce more consistent samples. Funding for these ongoing projects comes from the IME department, the PI's research funds, and from Genius Metal.

An extension of the project has also been approved to study the effects of coatings on the best performing tools from the tests. The PI and Dr. Liu will work to select an appropriate coating material and scheme for the tools. Genius Metal will also produce additional tool samples and arrange for the coatings of these new tools.

Lastly, the PI and Dr. Liu from Genius Metal have submitted a proposal in December 2006 to the Small Business Innovative Research (SBIR) grant funding track at NSF under the Emerging Opportunities/Manufacturing Materials topic. The SBIR proposal is for a 6-month Phase 1 effort to continue with the testing and development of the alternative-binder tooling materials that are being worked on for this project. If successful for Phase 1 (or Phase 2 which is a 2-year effort), the additional resources will only serve to improve the chances of success for the current development effort.

## **Crack diagnosis of a rotor from torsional/lateral coupling mechanism**

Project Investigators:

Xi Wu and James Meagher  
Mechanical Engineering  
California Polytechnic State University  
San Luis Obispo, California

## Project Overview

The purpose of this investigation into the vibration behavior of a rotor with transverse cracks is to determine whether and in which way a crack can be recognized by measuring the shaft vibrations so that catastrophic failures in rotors can be prevented and maintenance may be intelligently scheduled. Fatigue cracks/flaws are a potential source of catastrophic failures in rotating machinery. The sudden failures induced by fatigue cracks release considerable energy stored in the rotating system, sometimes causing a rotor to fly apart. Despite considerable efforts spent to develop a foolproof and reliable strategy to detect cracked rotors in aerospace, military, and power generation applications current strategies remain inadequate. ONR Science and Technology Objectives: Logistics Log STO-2: Predictive maintenance systems, and Maneuver MVR STO-1: Advanced power plants, drive trains, and suspensions relate to this study.

The long-term goal of the proposed research is to develop both analytical and experimental methods to quantitatively characterize the torsional/lateral responses of a two-mass cracked rotor. This study focuses on developing a nonlinear dynamic analytical model of a torsional/lateral coupled extended Jeffcott rotor with a crack, laterally loaded with a constant radial force and a rotating unbalance. This is considered to be an accurate yet tractable model that represents a wide class of machines. The contribution of this research is the development of a practical and realizable diagnostic method for rotating machinery which can predict the existence and severity of a cracked shaft from the vibration measurements typically available by high speed rotating machinery. The rotors of turbines and generators could be represented by a simplified equivalent configuration in which the masses of the turbine rotor and the generator rotor are concentrated on disks mounted on a light elastic shaft with two orthogonal lateral and two angular (torsional) degrees of freedom. Earlier work on the same topic simulated a crack as an asymmetric shaft, which is not accurate and cannot capture the whole profiles of a cracked rotor. Crack specific signatures predicted by simulation will be experimentally tested in future studies. As a result of mutual influence between lateral and torsional vibrations due to the crack, the vibration patterns, orbit shape and frequency spectra are unique and could be useful indicators for crack diagnosis. An expected outcome of the research is the development of more accurate and feasible nonlinear torsional/lateral coupled cracked model that may be used for automated machinery fault detection.

## Major Findings and Results

- A practical cracked rotor model representing a wide class of machines has been developed.
- Results of this model demonstrate several diagnostic indicators which may be useful to detect the presence and severity of shaft cracks:
  - lateral-torsional coupling.
  - torsional critical speed frequencies at fixed non-integer ratios of lateral to torsional natural frequency.
  - the spectrum of the torsional vibration is shown to contain super- synchronous critical speeds related to the lateral natural frequency.
  - for cracked rotors that are torsionally rigid rotors and without lateral vibration, substantial vibrations occur at one-third and one-half the fundamental torsional

resonance. With an increase in crack depth, the resonance peaks at  $1\times$ ,  $2\times$  and  $3\times$  of the torsional vibration increase dramatically.

- for more flexible rotors with a crack, strong coupling between lateral and torsional vibrations due to the crack and side load appear. Nonlinear lateral-torsional coupling from a crack also shifts the fundamental resonance peak in the torsional vibration response. The resonance peak frequencies shift depending on the ratio of the lateral to torsional natural frequencies with the peak responses occurring at fixed non-integer values of the lateral natural frequency for any particular frequency ratio.
- Results of this study have been accepted for presentation at the Society of Experimental Mechanics IMAC-XXV conference.
- We have leveraged this study to obtain hardware to make specialized measurements and increase our capabilities.
- Preliminary experimental measurements of lateral vibrations have been conducted in preparation for more extensive verification testing.
- Designed and built a machine allowing manufacture of cracked shafts for further study.

## Background

A comprehensive literature survey of various crack modeling techniques and system behaviors of cracked rotors can be found in Wauer [1] and Dimarogonas [2]. More than 300 papers are cited. Diagnostic indicators of cracked rotors reported in these cited studies include: changes in lateral resonances due to flexibility changes in a cracked rotor, coupling that can be recognized from additional harmonics in the frequency spectrum, sub-harmonic resonances at approximately half and one third of the bending critical speed of the rotor [3-5]. Nonlinear modeling of the crack began with simple step functions that were later improved upon by Mayes. Mayes model [6] is more practical for deep cracks than a hinged model. Based on Mayes modified model, Sawicki and Wu et al. [7, 8, 9] studied the transient vibration response of a cracked Jeffcott rotor under constant acceleration ratios and under constant external torque. Muszynska et al. [10] and Bently et al. [11] discuss rotor coupled lateral and torsional vibrations due to unbalance, as well as due to shaft anisotropy under a constant radial preload force. Their experimental results exhibited the existence of significant torsional vibrations due to coupling with the lateral modes. In Bently and Muszynska's experiments, an asymmetric shaft was used to simulate the behavior of a crack. Our research extends and is extensively based upon the investigations of both Bently et al. [11] and Wu's work [12]. Starting from energy equations, an analytical model with four degrees of freedom for a torsional/lateral coupled rotor due to a crack is developed. A radial constant force is applied to the outboard disk to emphasize the effects of the gravity force which plays a critical role for the "breathing" of a crack. As preload increases, the vibration amplitudes in both lateral and torsional directions increase, making the measurements much easier to measure. The "second-order" nonlinear coupling terms due to a crack introduce super-synchronous peaks at certain rotational speeds, which is unique for a cracked rotor and might be used as an unambiguous crack indicator. Computer simulations also show that the rotational speeds at which amplitudes of the torsional vibrations reach maximum are governed by the ratio of lateral to torsional natural frequency.

## Equations of Motion

An analytical model of an extended Jeffcott rotor, Figure 1, is first derived from Lagrange's equations taking into consideration lateral/torsional vibration coupling mechanism induced by a "breathing" of a crack. Four degrees of freedom describe the model: two lateral displacements and one torsional angular displacement of an outboard disk, and the torsional angular displacement of an inboard disk. The nonlinearities associated with a breathing crack couple the four equations of motion. The rotor is driven through a flexible coupling by an electric motor with a speed controller. A crack is located near the outboard disk. Shaft lateral vibrations are measured with a pair of proximity probes in horizontal and vertical orientations at the outboard disk where a downward constant radial force  $P$  is also applied. The coupled torsional-flexural vibrations are modeled using four degrees of freedom: torsional rotation at each disk and lateral motion at the outboard disk. Figure 2 show the section view of the cracked shaft, in both inertial  $(X, Y)$  and rotating coordinates  $(\xi, \eta)$ .

The angular position of the outboard disk is expressed as  $\Phi(t) = \Omega t + \varphi(t) - \varphi_0$ , where  $\Omega$  is a rotational speed of the rotor,  $\varphi(t)$  is a torsional angle of outboard disk relative to the motor, and  $\varphi_0$  is the initial angular position. Similarly, the angular position of the inboard disk is expressed as  $\Theta(t) = \Omega t + \theta(t) - \theta_0$ , where  $\theta(t)$  is the torsional angle of the inboard disk relative to the motor. The outboard disk's vibration is represented by the torsional coordinate  $\Phi(t)$  and two lateral displacements in inertial coordinates. The inboard disk's vibration is described by the angle  $\Theta(t)$ . The location of the center of mass of the outboard disk can be expressed as the following:

$$\begin{aligned} x_{cm} &= X + \varepsilon \cos(\Phi + \delta) \\ y_{cm} &= Y + \varepsilon \sin(\Phi + \delta) \end{aligned} \quad (1)$$

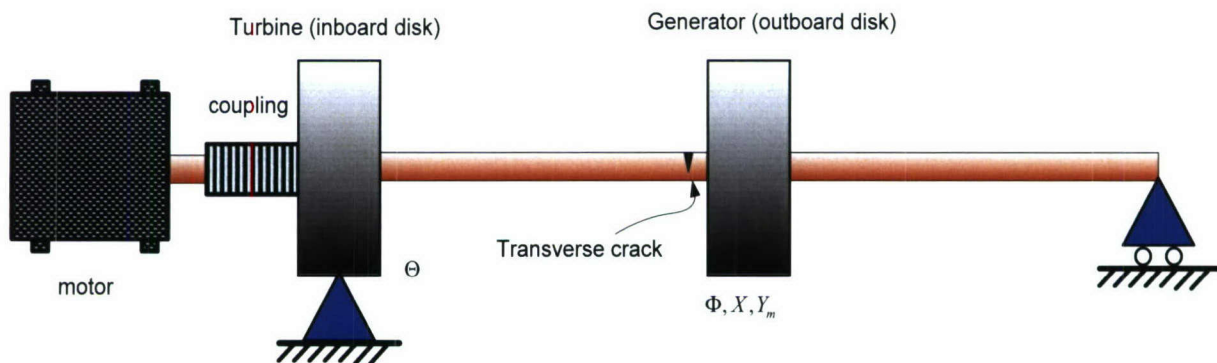
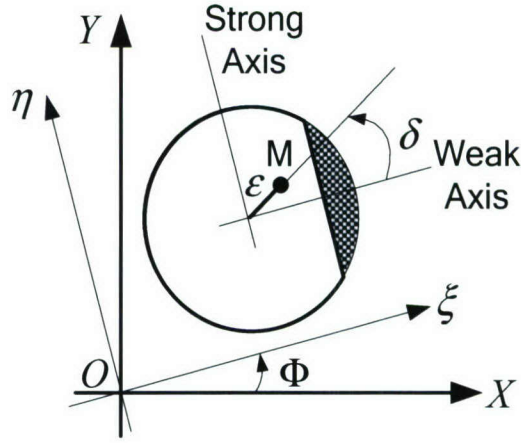


Figure 1. Configuration of the cracked extended Jeffcott rotor with two disks.



**Figure 2. Section view of cracked shaft.**

The kinetic energy, potential energy and dissipation function for the cracked rotor system can respectively be expressed as the following:

$$\begin{aligned}
 T &= \frac{1}{2} I (\Omega + \dot{\phi})^2 + \frac{1}{2} I_0 (\Omega + \dot{\theta})^2 + \frac{M}{2} \{ \dot{X}^2 + \dot{Y}^2 - 2\epsilon \dot{X} (\Omega + \dot{\phi}) \sin(\Phi + \delta) \\
 &\quad + 2\epsilon \dot{Y} (\Omega + \dot{\phi}) \cos(\Phi + \delta) + \epsilon^2 (\Omega + \dot{\phi})^2 \} \\
 U &= \frac{1}{2} (k_{11} X^2 + k_{22} Y^2) + k_{12} XY + \frac{1}{2} K_t (\phi - \theta)^2 \\
 (2) \quad D &= \frac{1}{2} C \dot{X}^2 + \frac{1}{2} C \dot{Y}^2 + \frac{1}{2} C_t (\dot{\theta} - \dot{\phi})^2
 \end{aligned}$$

The damping is modeled as lumped viscous damping at the outboard disk and lumped torsional viscous damping of the shaft. The stiffness matrix for a Jeffcott rotor with a cracked shaft in inertial coordinates,  $\mathbf{K}_I$  is given [4,7,8] by the following. Details can be found in reference [12].

$$\mathbf{K}_I = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} = \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix} - \frac{f(\Phi)K}{2} \begin{pmatrix} \Delta k_1 + \Delta k_2 \cos 2\Phi & \Delta k_2 \sin 2\Phi \\ \Delta k_2 \sin 2\Phi & \Delta k_1 - \Delta k_2 \cos 2\Phi \end{pmatrix} \quad (3)$$

$$\text{Where } \Delta k_1 = \frac{\Delta k_\xi + \Delta k_\eta}{K} \text{ and } \Delta k_2 = \frac{\Delta k_\xi - \Delta k_\eta}{K} \quad (4)$$

$\Delta k_\xi$ ,  $\Delta k_\eta$  are respectively the reduced stiffness in  $\xi$  and  $\eta$  directions in a rotor-fixed coordinate system.  $f(\Phi) = \frac{1 + \cos(\Phi)}{2}$  is a steering function which Mayes and Davies [6] proposed to

illustrate a smooth transition between the opening and closing of a “breathing” crack in rotating coordinates.  $\Delta k_\eta = \Delta k_\xi / 6$  is assumed to describe the stiffness variation for deep cracks.

The general equations of motion are obtained using Lagrange’s equations.

$$\begin{aligned} \ddot{X} + \frac{C}{M} \dot{X} + \omega_n^2 \left[ 1 - \frac{f(\Phi)}{2} (\Delta k_1 + \Delta k_2 \cos 2\Phi) \right] X - \frac{\omega_n^2 f(\Phi) \Delta k_2 \sin 2\Phi}{2} \left( Y_m - \frac{P}{K} \right) \\ = \varepsilon (\Omega + \dot{\phi})^2 \cos(\Phi + \delta) + \varepsilon \ddot{\phi} \sin(\Phi + \delta) \end{aligned}$$

$$\begin{aligned} \ddot{Y}_m + \frac{C}{M} \dot{Y}_m - \frac{\omega_n^2 f(\Phi) \Delta k_2 \sin 2\Phi}{2} X + \omega_n^2 \left[ 1 - \frac{f(\Phi)}{2} (\Delta k_1 - \Delta k_2 \cos 2\Phi) \right] Y_m \\ = \varepsilon (\Omega + \dot{\phi})^2 \sin(\Phi + \delta) - \varepsilon \ddot{\phi} \cos(\Phi + \delta) - \frac{Pf(\Phi)}{2M} (\Delta k_1 - \Delta k_2 \cos 2\Phi) \end{aligned} \quad (5)$$

$$\begin{aligned} \ddot{\theta} + \frac{K_t + K_c}{I_0} \theta - \frac{K_t}{I_0} \varphi = -\frac{C_t + C_c}{I_0} \dot{\theta} + \frac{C_t}{I_0} \dot{\varphi} \\ \ddot{\varphi} + \frac{C_t}{I} \dot{\varphi} - \frac{C_t}{I} \dot{\theta} + \frac{K_t}{I} \varphi - \frac{K_t}{I} \theta = \frac{P \varepsilon f(\Phi)}{2I} (\Delta k_1 \cos(\Phi + \delta) - \Delta k_2 \cos(\Phi - \delta)) \\ + \frac{P^2}{2KI} \left[ \frac{1}{2} \frac{\partial f(\Phi)}{\partial \Phi} (\Delta k_1 - \Delta k_2 \cos 2\Phi) + f(\Phi) \Delta k_2 \sin 2\Phi \right] + \Gamma_c \end{aligned}$$

Where:

$$\begin{aligned} \Gamma_c = -\frac{C\varepsilon}{\rho^2 M} \left[ \dot{X} \sin(\Phi + \delta) - \dot{Y}_m \cos(\Phi + \delta) \right] + \frac{\varepsilon \omega_n^2}{\rho^2} \left( 1 - \frac{f(\Phi)}{2} \Delta k_1 \right) \left[ -X \sin(\Phi + \delta) + Y_m \cos(\Phi + \delta) \right] \\ + \frac{\varepsilon f(\Phi) \Delta k_2 \omega_n^2}{2\rho^2} \left[ -X \sin(\Phi - \delta) + Y_m \cos(\Phi - \delta) \right] \\ + \frac{X^2 \omega_n^2}{2\rho^2} \left[ \frac{1}{2} \frac{\partial f(\Phi)}{\partial \Phi} (\Delta k_1 + \Delta k_2 \cos 2\Phi) - f(\Phi) \Delta k_2 \sin 2\Phi \right] \\ + \frac{Y_m \left( Y_m - \frac{2P}{K} \right) \omega_n^2}{2\rho^2} \left[ \frac{1}{2} \frac{\partial f(\Phi)}{\partial \Phi} (\Delta k_1 - \Delta k_2 \cos 2\Phi) + f(\Phi) \Delta k_2 \sin 2\Phi \right] \\ + \frac{\Delta k_2 X \left( Y_m - \frac{P}{K} \right) \omega_n^2}{2\rho^2} \left[ \frac{\partial f(\Phi)}{\partial \Phi} \sin 2\Phi + 2f(\Phi) \cos 2\Phi \right] \end{aligned} \quad (6)$$

$Y = Y_m - \frac{P}{K}$  is used to delineate the static offset from dynamic response. Also, since  $\varepsilon^2 \approx 0$ ,  $I = I_0 + \varepsilon^2 M \approx I_0$  is assumed. Using non-dimensionalized time defined by the following:

$$\tau = \omega_n t, \quad \frac{d(\cdot)}{dt} = \omega_n \frac{d(\cdot)}{d\tau} = \omega_n (\cdot)', \quad \frac{d^2(\cdot)}{dt^2} = \omega_n^2 \frac{d^2(\cdot)}{d\tau^2} = \omega_n^2 (\cdot)'' \quad (7)$$

equations (5) and (6) take the following form:

$$\begin{aligned} X'' + 2\zeta X' + \left[ 1 - \frac{f(\Phi)}{2} (\Delta k_1 + \Delta k_2 \cos 2\Phi) \right] X - \frac{f(\Phi) \Delta k_2 \sin 2\Phi}{2} \left( Y_m - \frac{P}{M\omega_n^2} \right) \\ = \varepsilon \left( \frac{\Omega}{\omega_n} + \varphi' \right)^2 \cos(\Phi + \delta) + \varepsilon \varphi'' \sin(\Phi + \delta) \\ Y_m'' + 2\zeta Y_m' - \frac{f(\Phi) \Delta k_2 \sin 2\Phi}{2} X + \left[ 1 - \frac{f(\Phi)}{2} (\Delta k_1 - \Delta k_2 \cos 2\Phi) \right] Y_m \\ = \varepsilon \left( \frac{\Omega}{\omega_n} + \varphi' \right)^2 \sin(\Phi + \delta) - \varepsilon \varphi'' \cos(\Phi + \delta) - \frac{P}{M} \frac{f(\Phi)}{2\omega_n^2} (\Delta k_1 - \Delta k_2 \cos 2\Phi) \end{aligned} \quad (8)$$

$$\begin{aligned} \theta'' + (1 + K_r) \left( \frac{\omega_t}{\omega_n} \right)^2 \theta - \left( \frac{\omega_t}{\omega_n} \right)^2 \varphi = -2\zeta_t (1 + C_r) \frac{\omega_t}{\omega_n} \theta' + 2\zeta_t \frac{\omega_t}{\omega_n} \varphi' \\ \varphi'' + 2\zeta_t \frac{\omega_t}{\omega_n} \varphi' - 2\zeta \frac{\omega_t}{\omega_n} \theta' + \left( \frac{\omega_t}{\omega_n} \right)^2 \varphi - \left( \frac{\omega_t}{\omega_n} \right)^2 \theta = \frac{P}{2M} \frac{\varepsilon f(\Phi)}{\omega_n^2 \rho^2} (\Delta k_1 \cos(\Phi + \delta) - \Delta k_2 \cos(\Phi - \delta)) \\ + \frac{P^2}{2M^2} \frac{1}{\omega_n^4 \rho^2} \left[ \frac{1}{2} \frac{\partial f(\Phi)}{\partial \Phi} (\Delta k_1 - \Delta k_2 \cos 2\Phi) + f(\Phi) \Delta k_2 \sin 2\Phi \right] + \frac{\Gamma_c}{\omega_n^2} \end{aligned}$$

$$\begin{aligned}
\frac{\Gamma_c}{\omega_n^2} = & -2\zeta \frac{\varepsilon}{\rho^2} [X' \sin(\Phi + \delta) - Y'_m \cos(\Phi + \delta)] + \frac{\varepsilon}{\rho^2} \left(1 - \frac{f(\Phi)}{2} \Delta k_1\right) [-X \sin(\Phi + \delta) + Y_m \cos(\Phi + \delta)] \\
& + \frac{\varepsilon f(\Phi) \Delta k_2}{2\rho^2} [-X \sin(\Phi - \delta) + Y_m \cos(\Phi - \delta)] + \frac{X^2}{2\rho^2} \left[ \frac{1}{2} \frac{\partial f(\Phi)}{\partial \Phi} (\Delta k_1 + \Delta k_2 \cos 2\Phi) - f(\Phi) \Delta k_2 \sin 2\Phi \right] \\
& + \frac{Y_m \left( Y_m - \frac{2P}{M\omega_n^2} \right)}{2\rho^2} \left[ \frac{1}{2} \frac{\partial f(\Phi)}{\partial \Phi} (\Delta k_1 - \Delta k_2 \cos 2\Phi) + f(\Phi) \Delta k_2 \sin 2\Phi \right] \\
& + \frac{\Delta k_2 X \left( Y_m - \frac{P}{M\omega_n^2} \right)}{2\rho^2} \left[ \frac{\partial f(\Phi)}{\partial \Phi} \sin 2\Phi + 2f(\Phi) \cos 2\Phi \right]
\end{aligned} \tag{9}$$

Where:  $K_c = K_r K_t$ ,  $C_c = C_r C_t$

Two special cases can be derived from the above general case.

Case 1 (pure torsional vibration)

Assuming no lateral vibration,  $X = 0$ ,  $Y_m = 0$ , and a rigid drive coupling,  $\dot{\Theta} = \Omega$ , leads to the following simplification:

$$\begin{aligned}
\ddot{\phi} + \frac{C_t}{I} \dot{\phi} + \frac{K_t}{I} \phi = & \frac{P\varepsilon}{8I} (\Delta k_1 - \Delta k_2) + \left\{ \frac{P\varepsilon}{4I} (\Delta k_1 - \Delta k_2) \cos \Phi + \frac{P^2}{8KI} \left( -\Delta k_1 + \frac{\Delta k_2}{2} \right) \sin \Phi \right\} \\
& + \left\{ \frac{P\varepsilon}{8I} (\Delta k_1 - \Delta k_2) \cos 2\Phi + \frac{P^2 \Delta k_2}{4KI} \sin 2\Phi \right\} + \frac{3P^2 \Delta k_2}{16KI} \sin 3\Phi
\end{aligned} \tag{10}$$

Introducing the following two constants:

$$E_1 = \frac{P\varepsilon}{2I\omega_n^2} = \frac{P}{2M} \frac{\varepsilon}{\rho^2 \omega_n^2}, \quad E_2 = \frac{P^2}{2KI\omega_n^2} = \frac{P^2}{2M^2} \frac{1}{\rho^2 \omega_n^4} \tag{11}$$

The non-dimensional form of equation (10) becomes:

$$\begin{aligned}
\phi'' + 2\zeta_t \frac{\omega_t}{\omega_n} \phi' + \left( \frac{\omega_t}{\omega_n} \right)^2 \phi = & \frac{E_1}{4} (\Delta k_1 - \Delta k_2) + \left\{ \frac{E_1}{2} (\Delta k_1 - \Delta k_2) \cos \Phi + \frac{E_2}{4} \left( -\Delta k_1 + \frac{\Delta k_2}{2} \right) \sin \Phi \right\} \\
& + \left\{ \frac{E_1}{4} (\Delta k_1 - \Delta k_2) \cos 2\Phi + \frac{E_2 \Delta k_2}{2} \sin 2\Phi \right\} + \frac{3E_2 \Delta k_2}{8} \sin 3\Phi
\end{aligned} \tag{12}$$

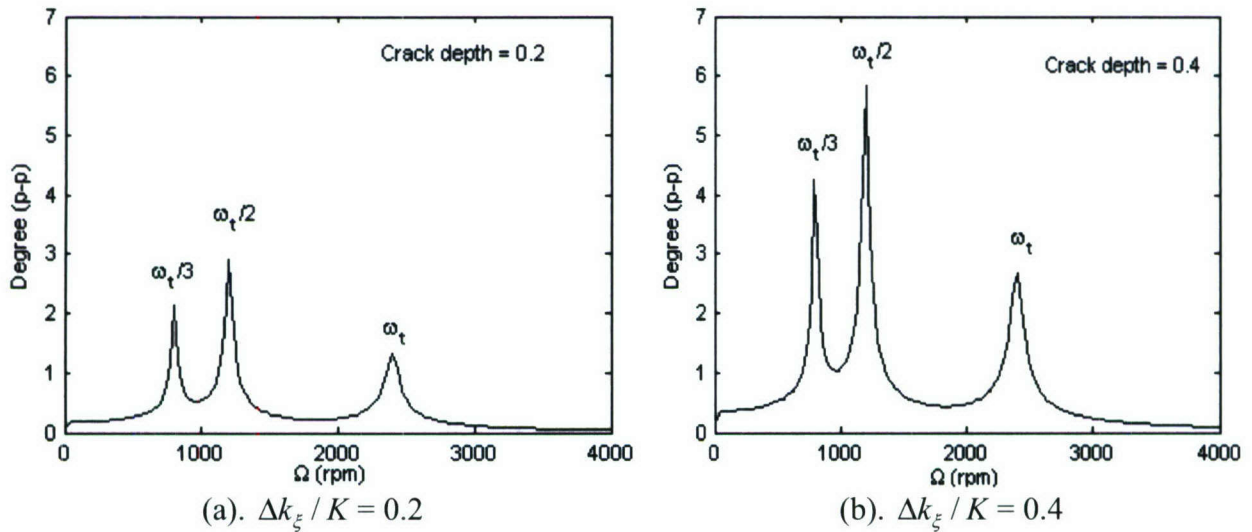
Case 2 (lateral vibration without torsional vibration)

In this case the cracked rotor is considered as laterally flexible and torsionally rigid. Therefore, the system could be described by the first two equations of (5) in which  $\varphi$  and  $\theta$  are assumed zero. Since similar scenarios are well studied in [4,5,7,12], it is not explored here except to note that the general model presented in this work reduces to known models under these assumptions.

$$\begin{aligned}
 & \ddot{X} + \frac{C}{M} \dot{X} + \omega_n^2 \left[ 1 - \frac{f(\Omega t)}{2} (\Delta k_1 + \Delta k_2 \cos(2\Omega t)) \right] X - \frac{\omega_n^2 f(\Omega t) \Delta k_2 \sin(2\Omega t)}{2} \left( Y_m - \frac{P}{K} \right) \\
 &= \varepsilon \Omega^2 \cos(\Omega t + \delta) \\
 & \ddot{Y}_m + \frac{C}{M} \dot{Y}_m - \frac{\omega_n^2 f(\Omega t) \Delta k_2 \sin(2\Omega t)}{2} X + \omega_n^2 \left[ 1 - \frac{f(\Omega t)}{2} (\Delta k_1 - \Delta k_2 \cos(2\Omega t)) \right] Y_m \\
 &= \varepsilon \Omega^2 \sin(\Omega t + \delta) - \frac{Pf(\Omega t)}{2M} (\Delta k_1 - \Delta k_2 \cos(2\Omega t))
 \end{aligned}
 \tag{13}$$

### Simulation Results

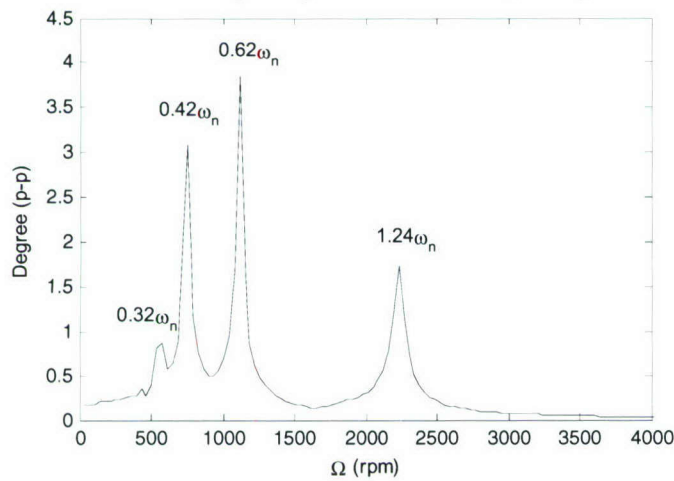
Computer simulation results using the parameters in table 1 for the case of pure torsion without lateral vibration are illustrated in figure 3. This response which is calculated using equation 10 can be interpreted as a nonlinear oscillator with 1X excitation  $\left\{ \frac{P\varepsilon}{4I} (\Delta k_1 - \Delta k_2) \cos \Phi + \frac{P^2}{8KI} \left( -\Delta k_1 + \frac{\Delta k_2}{2} \right) \sin \Phi \right\}$ , 2X excitation  $\left\{ \frac{P\varepsilon}{8I} (\Delta k_1 - \Delta k_2) \cos 2\Phi + \frac{P^2 \Delta k_2}{4KI} \sin 2\Phi \right\}$ , and a 3X excitation  $\left\{ \frac{3P^2 \Delta k_2}{16KI} \sin 3\Phi \right\}$  due to the unbalance, the depth of the crack and the side load.



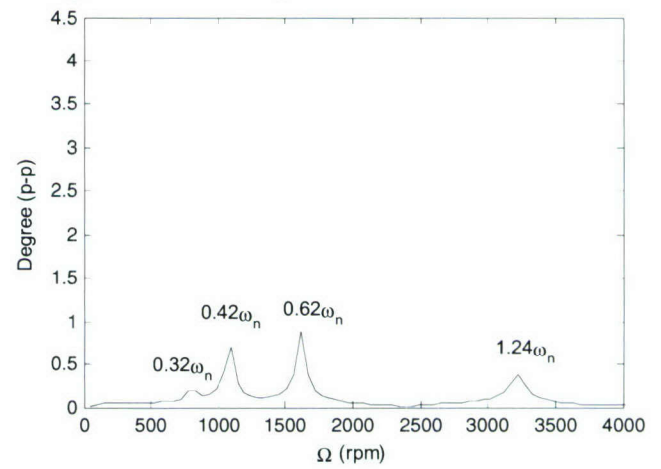
**Figure 3. Peak-to-peak torsional vibration response,  $\phi$ , from case 1.**

In addition to the excitations noted above, the constant forcing excitation  $\frac{P\epsilon}{8I}(\Delta k_1 - \Delta k_2)$  is also present. It is worthy to notice that the crack depths  $\Delta k_\xi$ ,  $\Delta k_\eta$ , contained in  $\Delta k_1, \Delta k_2$  which are originally introduced from the lateral stiffness terms, significantly affect the torsional vibration amplitude. Note that these vibrations would require torsional transducers for measurement.

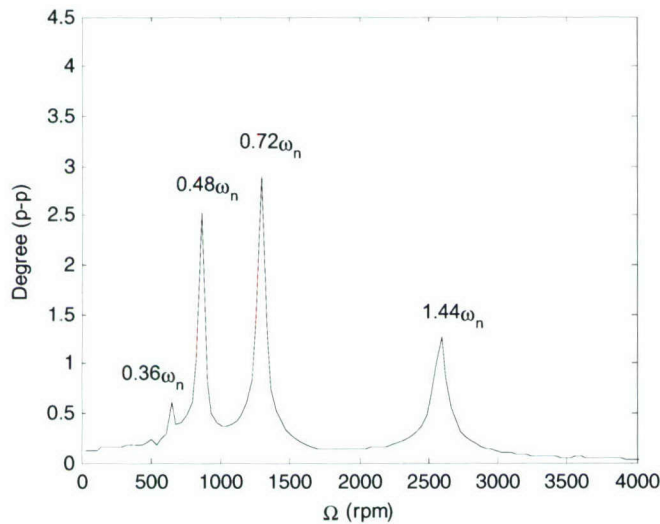
When the parameters shown in table 2 are used in the general four degree of freedom model, the critical frequencies shift from those shown above (see figure 4). Critical speeds are no longer at integer fractional multiples of the torsional natural frequencies. Instead, for a given ratio of torsional to lateral natural frequency critical speeds occur at fixed non-integer multiples of the lateral natural frequency. When the frequency ratio is changed the critical speeds shift.



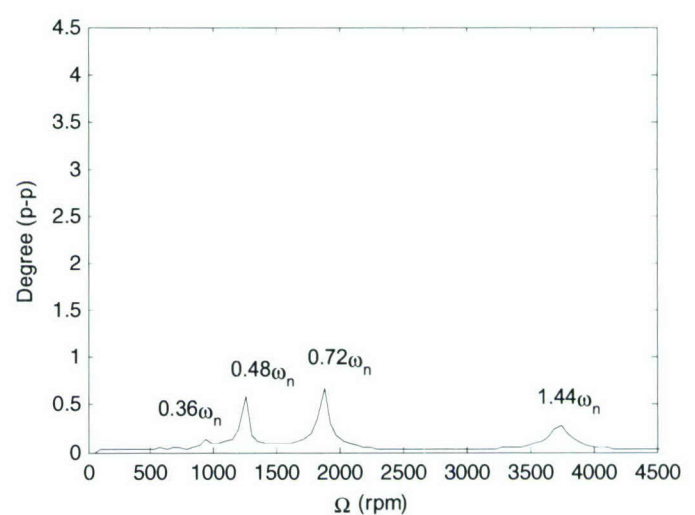
(a).  $\omega_t / \omega_n = 1.38$ ;  $\omega_n = 1800$  rpm



(b).  $\omega_t / \omega_n = 1.38$ ;  $\omega_n = 2600$  rpm



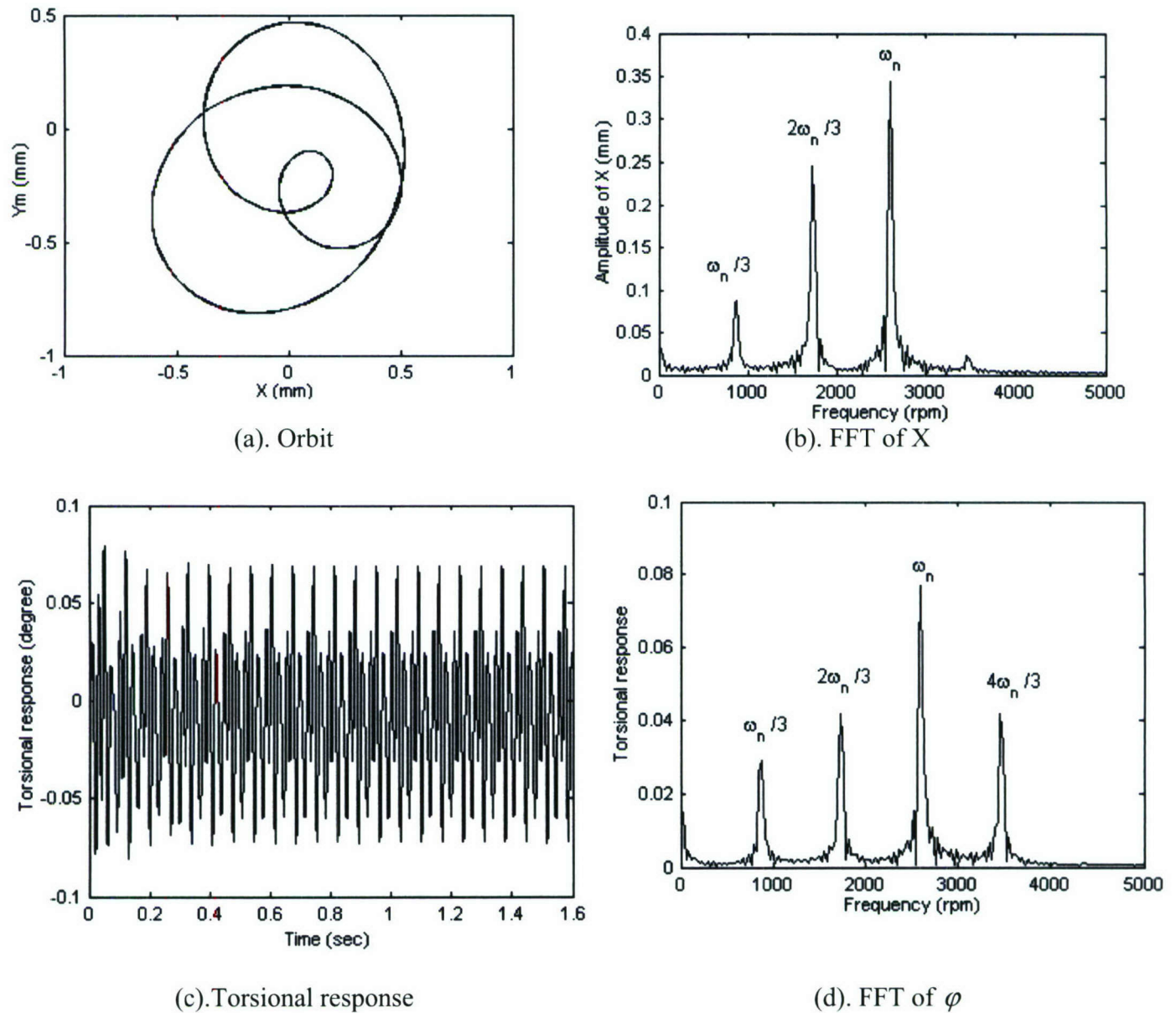
(c).  $\omega_t / \omega_n = 1.6$ ;  $\omega_n = 1800$  rpm



(d).  $\omega_t / \omega_n = 1.6$ ;  $\omega_n = 2600$  rpm

**Figure 4. Peak-to-peak torsional vibration response  $\phi$  from general lateral/torsional coupling case.**

The steady state response at one third the lateral natural using the same parameters as in (d) above is shown in figure 5. The orbit has three loops from the  $3x$  excitation. The x-probe shows significant super-synchronous response. The spectrum of the torsional vibration contains super-synchronous critical speeds related to the lateral natural frequency.



**Figure 5. The vibration responses at  $\Omega = \omega_n/3$  ;  $\omega_t/\omega_n = 1.38$ ;  $\omega_n = 2600$  rpm.**

**Table 1: Model Physical Parameters for Pure Torsional Vibration, Case 1**

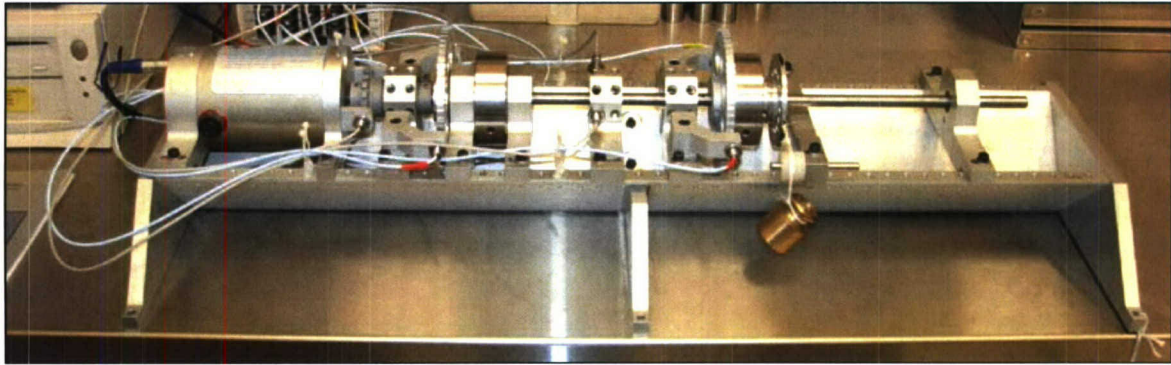
Parameters	Value	Units
$\omega_n$	10800	rpm
$\omega_t$	2400	rpm
$\varepsilon$	$15.24 \times 10^{-5}$	m
$\rho$	0.02286	m
$P/M$	1016	$\text{m/s}^2$
$\zeta_t$	0.02	

**Table 2: Model Physical Parameters for Torsional and Lateral Vibration, General Case**

Parameters	Value	Units
$\varepsilon$	$5.08 \times 10^{-5}$	m
$\rho$	0.02286	m
$P/M$	101.6	$\text{m/s}^2$
$\zeta_t$	0.02	
$\zeta$	0.1	
$K_r$	5	
$C_r$	1	
$\Delta k_{\xi} / K$	0.38	

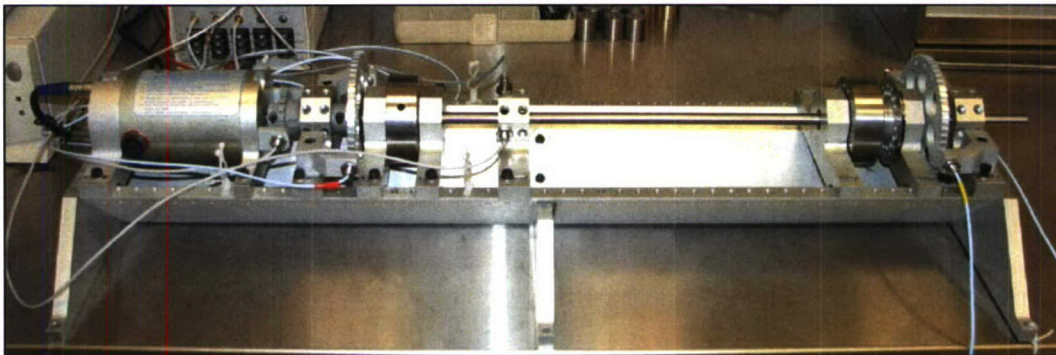
### Experimental Studies

An experiment is currently under development to simulate a real machine with a cracked shaft. The objectives of the experiment are to provide information for verifying and improving the mathematical model, and to investigate the practical implications of lateral and torsional vibration measurement as diagnostic tools. To this aim, the experimental foundation is the Bently Nevada RK-4 Rotor Kit shown in Figure 6. The assembly allows a wide range of configurations and is adaptable for the specific needs of this project. Two disks fixed to a common shaft and supported with plain rigid bearings are driven by an electric motor through a flexible coupling. A shaft-mounted bearing allows the application of a side load. The inboard disk is constrained by adjacent bearings to prevent lateral vibration. The outboard disk is supported by an adjustable bearing span to allow lateral flexibility. Torsional flexibility arises from the section of shaft connecting the two disks. Proximity probes are used to measure vibration and allow feedback control of rotor speed.



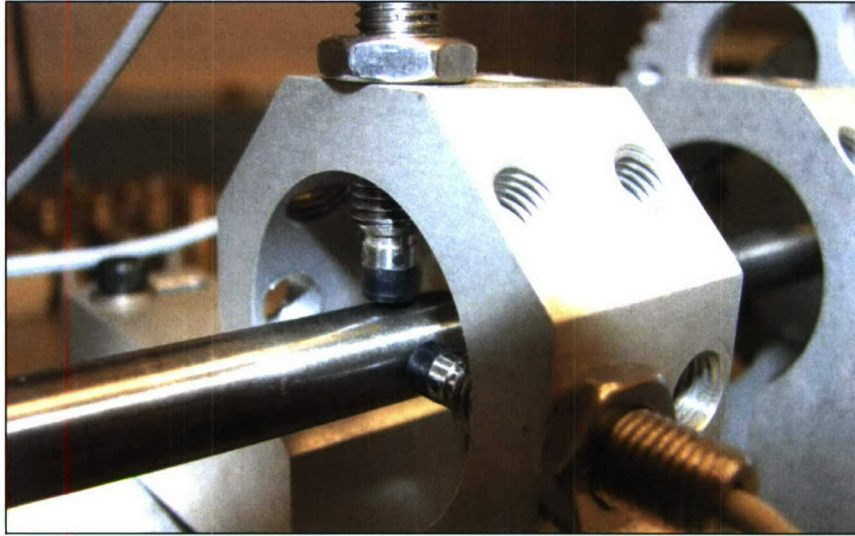
**Figure 6. Bently Nevada RK-4 Rotor Kit configured for a side load rotor with a cracked shaft**

**Configurability.** Based on the size of disks and shafts used by the RK-4 rotor kit, the bearing span of the outboard mass can be configured to provide a lateral natural frequency as low as 2,500rpm or a torsional natural frequency as low as 5,700rpm. Figure 6 shows the rotor configured for minimum lateral natural frequency while Figure 7 shows a possible configuration for minimizing the torsional natural frequency. Additional disks can be added to further reduce both natural frequencies.



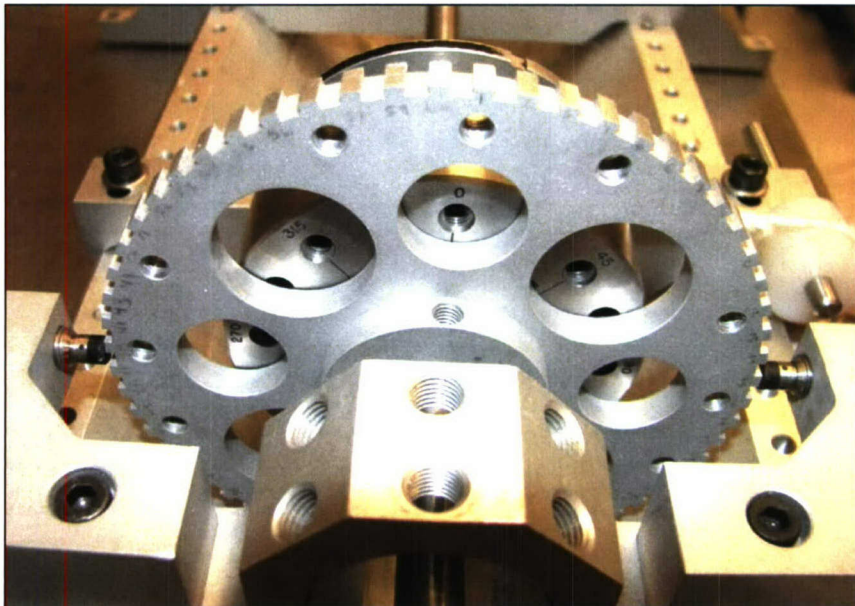
**Figure 7. RK-4 Rotor Kit configured for minimum torsional natural frequency**

**Lateral Measurement.** Lateral vibration is measured directly at the shaft using eddy current proximity probes as shown in Figure 8. The probes are oriented mutually perpendicular to measure vertical and horizontal displacement respectively.



**Figure 8. Proximity probes for lateral vibration measurement**

**Torsional Measurement.** Each disk is equipped with a transducer for measuring torsional vibration. The primary components are a custom machined gear wheel attached to the shaft and two proximity probes that are fixed 180 degrees apart in the horizontal plane, as shown in Figure 9.



**Figure 9. Torsional transducer consisting of a gear wheel and two proximity probes**

The principle of operation of the torsional transducer can be described as follows. As the gear wheel rotates, proximity probes detect the passage of teeth. The time it takes for teeth to pass can be attributed to either lateral or torsional motion as follows:

$$\Delta t_L = \frac{2\pi R}{N(R\Omega_i + V_y)} \text{ and } \Delta t_R = \frac{2\pi R}{N(R\Omega_i - V_y)} \quad (14)$$

where  $\Omega_i$  is the instantaneous rotative speed,  $V_y$  is the vertical velocity,  $R$  is the radius of the gear wheel,  $N$  is the number of teeth, and  $\Delta t_L$  and  $\Delta t_R$  are the times between pulses generated by the left and right probes respectively. The lateral component of the vibration signal ( $V_y$ ) can be eliminated from the above equations and the instantaneous rotative speed can be expressed as

$$\Omega_i = \frac{\pi}{N} \left( \frac{\Delta t_L + \Delta t_R}{\Delta t_L \Delta t_R} \right) \quad (15)$$

In general, the vibration of the disk can be express as

$$\theta = (\Omega_{motor} - \Omega_i)t \quad (16)$$

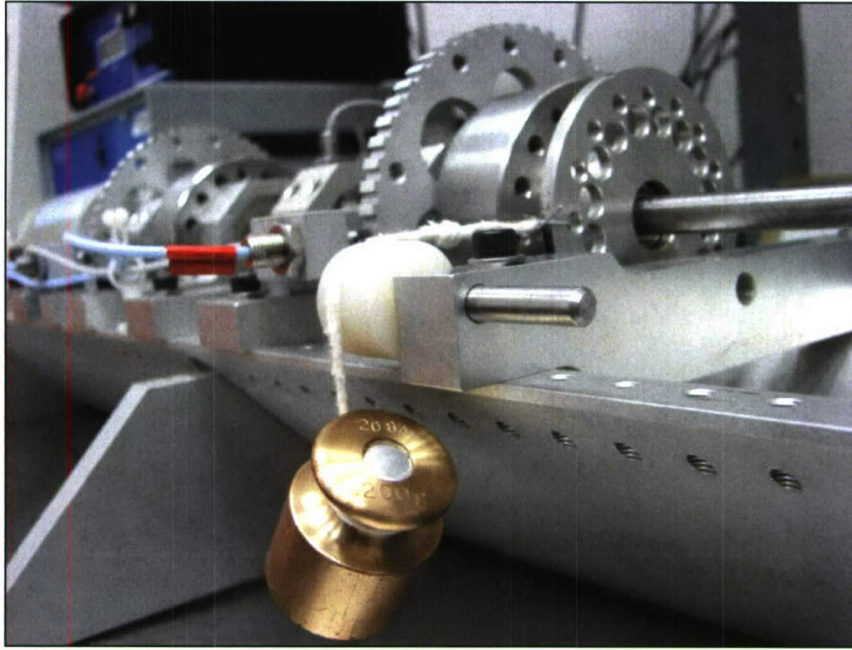
where  $\theta$  is the general vibration angle with respect to the motor and  $\Omega_{motor}$  is the rotative speed of the motor.



**Figure 10. Torsional vibration signal conditioner**

To simplify measurement of torsional vibration and reduce post processing, Bently Nevada TK-17 Torsional Vibration Signal Conditions are employed, as shown in Figure 10. These signal conditioners work from the same principles described above to process the output of the two proximity probes into a single output proportional to the torsional vibration amplitude.

**Side Loading.** A bearing and pulley system are employed to provide an adjustable side load as shown in Figure 11. Calibrated masses are hung from a string which draws over a pulley and loads a shaft-mounted bearing near the disk. Additional weights can be added to change loading conditions.



**Figure 11. Side load provided by gravity via a pulley and bearing system**

**Control and Data Acquisition.** The RK-4 Rotor Kit includes the speed controller shown in Figure 12. Motor speed can be maintained steady or ramped at an adjustable rate.



**Figure 12. RK-4 speed controller**

Two data acquisition systems are employed. First, the Bently Nevada ADRE 208-P (shown in Figure 13) works in unison with a personal computer to store and process data in real time, allowing for vibration monitoring and limited data storage.

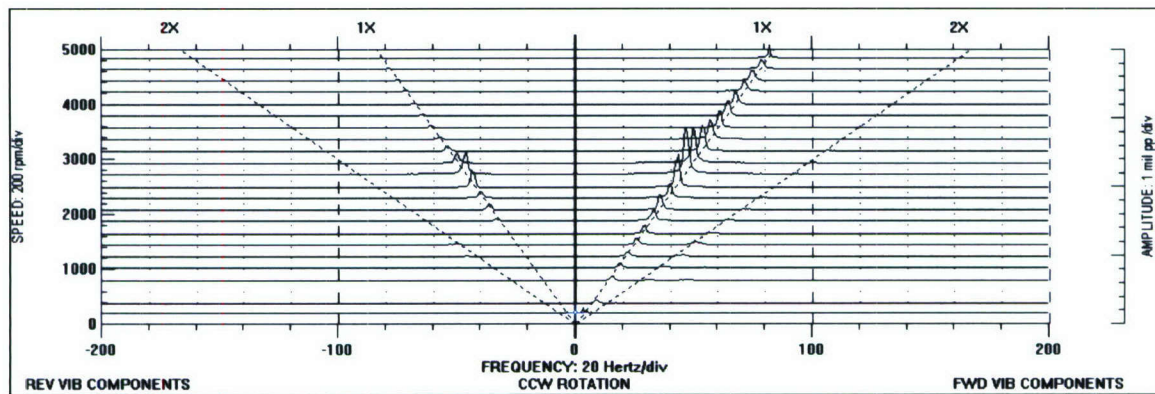


**Figure 13. Bently Nevada ADRE 208-P data acquisition system**

A second computer (not shown) records vibration data via a National Instruments SC-2345 Signal Conditioner and data acquisition adaptor.

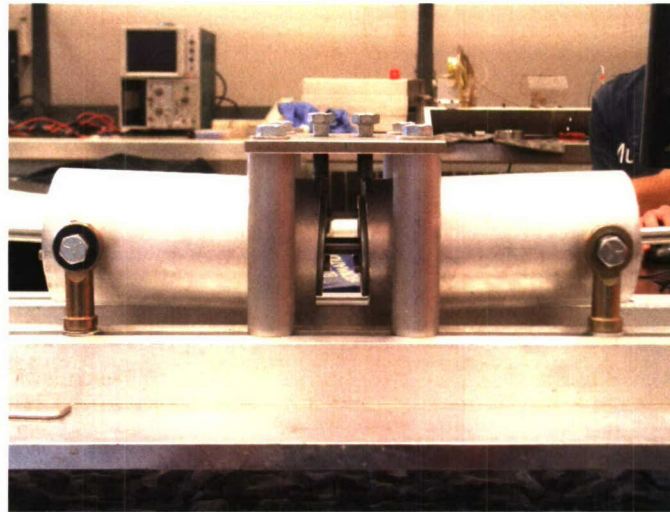
**Experimental Results.** Most of the hardware has recently been acquired for the experiment. Preliminary testing of an asymmetric shaft in the configuration shown in 6 have begun. An asymmetry was produced by grinding lateral notches near the disk. Stiffness of the shaft before and after notching was found experimentally by installing the shaft into bearings, hanging calibrated masses, and measuring deflection with a dial indicator. The full spectrum cascade plot of a representative trial is shown in

Figure 14. A small 2x forward vibration component is observed at a spin speed of half the lateral natural frequency, which may be indicative of the shaft asymmetry.



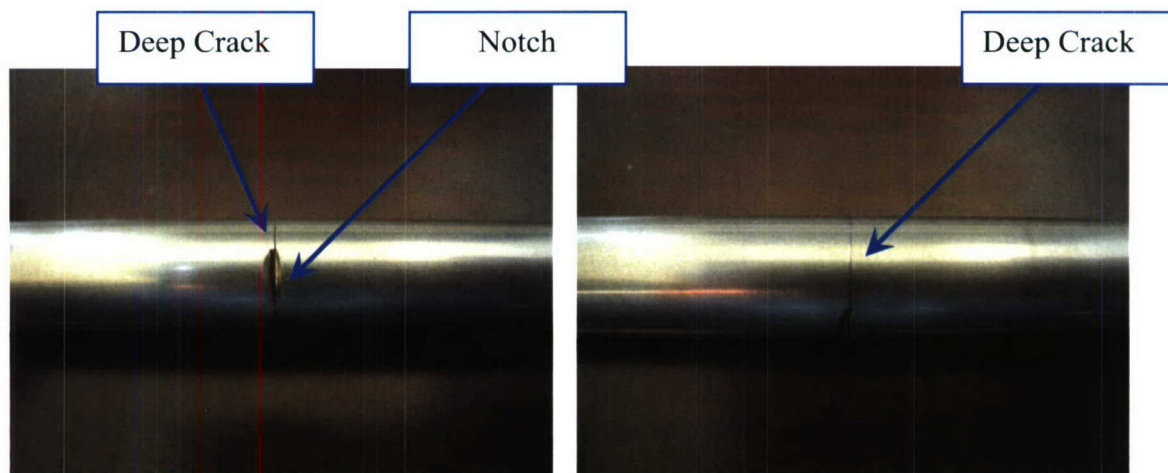
**Figure 14. Full spectrum cascade plot for a laterally loaded rotor**

After testing methodology is refined, shafts with manufactured transverse cracks will be used to test the theory developed in this project. An apparatus based on a rotating beam fatigue machine was built for use in our lab and is shown in Figure 15. This device places a shaft under cyclic bending by rotating it within side-loaded self-aligning bearings. 10 mm. AISI 440 Stainless Steel shafts produce propagating cracks after approximately 2 hours of operation.



**Figure 15. Shaft-crack manufacture**

One of the shafts created thus far is shown below in Figure 16. A notch is used to create a stress concentration and control where a crack will initiate. So far shafts with approximately 10% and 40% through cracks have been produced. A range of shafts with various crack depths and axial locations will be manufactured for use in future studies.



**Figure 16. Deep Crack created through the notching area**

### Future Goals.

- Complete development of and introduce torsional measurement.
- Perform parametric studies varying side load, imbalance, and configuration.
- Identify key parameters for tuning the mathematical simulation
- Verify, refute, or qualify findings of mathematical simulation
- Identify deficiencies and possible improvements in the mathematical model.

### Nomenclature

$\omega_n$ ,  $\omega_t$  lateral and torsional natural frequencies, respectively.

$\Theta$	angular displacement of inboard disk.
$\theta$	angular rotation of the inboard disk relative to motor.
$\theta_0$	initial angular location of inboard disk.
$\Phi$	angular rotation of outboard disk.
$\varphi$	angular rotation of the outboard disk relative to motor.
$\varphi_0$	initial angular location of outboard disk.
$I_0$	inboard disk polar moment of inertia.
$I$	outboard disk polar moment of inertia.
$\delta$	angular orientation of outboard disk eccentricity.
$\Gamma_c$	lateral coupling terms in outboard disk equation of motion.
$M$	outboard disk mass.
$P$	vertical side load.
$X, Y$	outboard disk lateral motion in inertial coordinates.
$\xi, \eta$	rotor fixed rotating coordinates.
$Y_m$	dynamic vertical vibration in inertial coordinates.
$\varepsilon$	eccentricity of outboard disk.
$C_c, K_c$	motor-shaft coupling damping and stiffness, respectively.
$C, C_t$	lateral and torsional damping coefficients, respectively.
$\zeta, \zeta_t$	lateral and torsional damping ratios, respectively.
$\Omega$	motor speed.
$K_t$	torsional shaft stiffness.
$K$	uncracked shaft lateral stiffness.
$f(\Phi)$	crack steering function.
$\rho$	radius of gyration.
$\Delta k_\xi, \Delta k_\eta$	the reduced stiffness in $\xi$ and $\eta$ directions, respectively.

## References

- [1]. Wauer, J., 1990, "On the Dynamics of Cracked Rotors: A Literature Survey", *Applied Mechanics Reviews*, Vol. 43(1), pp. 13-17.
- [2]. Dimarogonas, A. D., 1996, "Vibration of Cracked Structures: A State of the Art Review," *Engineering Fracture Mechanics*, Vol. 55 (5), pp. 831 – 857.
- [3]. Gasch, R., 1976, "Dynamic behavior of a simple rotor with a cross-sectional crack", *Paper C178/76, I. Mech. E. Conference on Vibrations in Rotating Machinery*, pp.123-128.
- [4]. Gasch, R. A, 1993, "Survey of the Dynamic Behavior of a Simple Rotating Shaft with a Transverse Crack", *Journal of Sound and Vibration* Vol. 162, pp313-332.
- [5]. Chan, R. K. C. and Lai, T. C., 1995, "Digital Simulation of a rotating shaft with a transverse Crack," *Appl. Math. Modelling* Vol. 19, July pp. 411-420.
- [6]. Mayes, I. W. and Davies, W. G. R., 1984, "Analysis of the Response of a Multi-Rotor-Bearing System Containing a Transverse Crack in a Rotor", *ASME Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 106, pp 139-145.

- [7]. Sawicki, J. T., Wu, X., Baaklini, G.Y. and Gyekenyesi, A., 2003, "Vibration-Based Crack Diagnosis in Rotating Shafts During Acceleration Through Resonance", *Proceedings of SPIE*, Vol. 5046.
- [8]. Sawicki, J. T., Bently, D. E., Wu, X., Baaklini, G.Y. and Friswell, M. I., 2003, "Dynamic Behavior of Cracked Flexible Rotor Subjected to Constant Driving Torque", ISCORMA-2, Gdańsk, Poland, 4-8, August 2003, pp. 231-241.
- [9]. Sawicki, J.T., Wu, X., Gyekenyesi, A. L. and Baaklini, G. Y., 2005 "Application of Nonlinear Dynamic Analysis for Diagnosis of Cracked Rotor Vibration Signatures", submitted to SPIE International Symposium, San Diego, California USA, March 7-10.
- [10]. Muszynska, A., Goldman, P. and Bently, D. E., 1992, "Torsional/Lateral Cross-Coupled Responses Due to Shaft Anisotropy: A New Tool in Shaft Crack Detection", *I. Mech. E.*, C432-090, Bath, United Kingdom, pp. 257-262.
- [11]. Bently, D. E., Goldman, P. and Muszynska, A., 1997, "“Snapping” Torsional Response of an Anisotropic Radially Loaded Rotor," *Journal of Engineering for Gas Turbines and Power*, Vol. 119, pp. 397-403.
- [12]. Wu, X., 2005, Doctoral dissertation, "Vibration-based Crack-induced Damage Detection of Shaft-disk System", Cleveland State University.

## **Publications and Presentations Related to Project**

### **Abstracts**

*Wu, X. and Meagher, J.; 'Investigation of coupled lateral and torsional vibrations of a cracked rotor under radial load'; Presentation to Society for Experimental Mechanics, Orlando, Florida USA, February 19 –22, 2007.*

A practical cracked rotor model with two disks representing a turbine and generator is studied using a four degree of freedom model; two transverse displacements and two torsional angular displacements. The differential equations of a rotor with a crack, unbalance, and constant radial force are first derived in detail using energy principles. The nonlinearities related with a "breathing" crack are incorporated, which distinguishes this paper from others studying stiffness anisotropy of a rotor with a similar rotor configuration. Through numerical simulations, this paper demonstrates the lateral-torsional coupling that occurs with a shaft crack and predicts torsional critical speed frequencies at fixed non-integer ratios of lateral to torsional natural frequency. The spectrum of the torsional vibration is shown to contain super-synchronous critical speeds related to the lateral natural frequency. The unique frequency response of the torsional motion predicted by this model could be employed for early detection of a cracked rotor.

## **Adaptive controller design with artificial neural networks**

Project Investigator:

Xiao-Hua (Helen) Yu  
Electrical Engineering  
California Polytechnic State University  
San Luis Obispo, California

## **Abstract**

In this research, a multi-layer feed-forward neural network controller is investigated for a class of Phase-Shifted Full-Bridge DC-DC converter. DC-DC converters can be found in almost every power electronics device. To improve the controller response to dynamical system changes, neural network has been chosen as an alternative to classic methods. However, no prior work has been done in the neural network approach to control a PSFB (Phase-Shifted Full-Bridge) converter yet. The neural network based controller proposed here has the advantage of adaptive learning ability, and can work under the situation when the input voltage and load current fluctuate. Levenberg-Marquardt back-propagation training algorithm is used in computer simulation and satisfactory experimental results are obtained.

## **1. Introduction**

DC-DC converters can be found in almost every electronic device nowadays, since all semiconductor components are powered by DC source. One of the design targets for electronic engineers is to improve the efficiency of power conversion. For PWM (pulse-width modulation) converters, switching loss is an important performance measure. Many different kinds of topologies ([1], [2], [3]) have been investigated in the past to reduce the switching loss. However, those topologies either need additional components for the power circuit, which may introduce some unstable factors to the circuit; or operate at variable frequency, which makes the filter design at output stage very difficult. Phase-shifted zero-voltage switching full-bridge converters overcome the above problems and thus have been received more and more attention recently. It employs zero-voltage-switching (ZVS) technique and allows the voltages across the transistors to swing to zero just before the start of the next conduction cycle ([4], [5], [6]).

Normally, the supply voltage and load current have a wide range of variation; so the controller has to be designed to work under such conditions. The conventional approach is to assume that the converter is operated around its equilibrium state; then a set of linear equations are derived based on this assumption [9] [10]. Artificial neural networks (ANN) have been widely used in the field of system identification, adaptive control, and statistical modeling in recent years. A neural network is composed of many non-linear adaptive processing elements and is capable of approximating any measurable function under certain conditions.

To improve the controller response of a DC-DC converter to dynamical system changes, neural network has been chosen as an alternative to classic methods ([7], [8]). However, no prior work has been done in the neural network approach for control of a PSFB (Phase-Shifted Full-Bridge) converter yet. In this research, a multi-layer feed-forward neural network controller is proposed. A Matlab Simulink model is developed first to generate the data set; then the neural network is trained by Levenberg-Marquardt back-propagation algorithm. Satisfactory simulation results are obtained.

## 2. Circuit Analysis

The circuit diagram of a PSFB DC-DC converter is shown in Fig. 1.

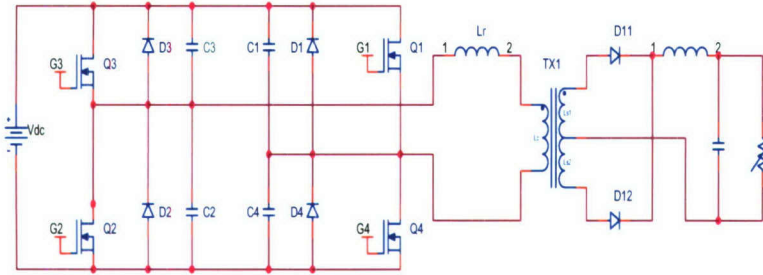


Fig 1. PSFB DC-DC Converter

Based on circuit analysis, the control scheme to drive the switching MOSFET (Q1, Q2, Q3 and Q4) is very complicated [6]. Based on circuit analysis, the control scheme to drive the switching MOSFET (Q1, Q2, Q3 and Q4) is very complicated [6]. The circuit is operated in one of the following modes:

**Mode 1:** The diagonal MOSFET Q3 & Q4 are conducting and power is delivered through the transformer to the load. The primary load current is flowing through the leakage inductance of the transformer. The total primary current is equal to the load current plus increasing magnetizing current of the transformer.

**Mode 2:** With Q4 on, the capacitance across Q1 is charged to +V. When Q4 turns OFF the current through the transformer inductance starts to charge the drain source capacitance of Q4, while at the same time discharges the capacitance of Q1. This action continues until the body diode of Q1 turns ON to clamp the voltage across Q1 at approximately 0.7V. The current through the transformer is sustained in the upper half of the power circuit.

**Mode 3:** When the voltage across Q1 reaches approximately 0V, Q1 turns ON. The time that is required for the capacitance of Q4 and Q1 to reach the desired voltage is determined by the characteristics of hardware. The current in this mode circulates through the conduction channels of Q3 and Q1.

**Mode 4:** Q3 turns off. The transformer current now starts to charge and discharge the capacitance of Q3 and Q2 respectively. It again requires a finite amount of time for the drain voltage of Q2 to reach "0 volt" at which point Q2 is allowed to be switched on. That means non-dissipate turn ON switching is accomplished.

**Mode 5:** With the complete discharge of its drain source capacitance Q2 is now ready to turn ON. Power is delivered to the load through the conduction path of Q1 and Q2 for an amount of time that is determined by the control circuit. Twice the product of this time, times the operating frequency of the oscillator gives the duty cycle of the converter as in a regular PWM converter. Duty Cycle  $D = 2 \cdot T_{on} \cdot F_s$ .

**Mode 6:** Following the power transfer by the above diagonal pair, Q1 turns OFF. The voltage across Q4 starts to decrease, and when this voltage reaches 0 volts the next mode starts.

**Mode 7:** In this mode, Q4 turns on and primary current circulates in the conduction channels of the lower pair.

**Mode 8:** Q2 turns OFF and current starts to charge and discharge the capacitance of Q2 and Q3 respectively. When the voltage across Q3 has reached 0 volts then Q3 turns ON non-dissipative and the complete cycle repeats itself from Mode 1.

The relationship between the input/output voltage and the duty cycle can be described as:

$$V_0 = \left( \frac{2nV_{in}}{T_s} \right) \left( \Phi \cdot \frac{T_s}{2} - T_1 - \frac{T_2}{2} \right) \quad (1)$$

where  $V_0$  is the output voltage,  $V_{in}$  is the input voltage, both in rms value;  $n$  is the transformer ratio,  $T_s$  is the switching period constant, and  $\Phi$  is the duty cycle.  $T_1$  and  $T_2$  can be obtained from the following equations:

$$T_1 = \frac{(nI + I_c)L_r}{V_{in}} \quad (2)$$

$$T_2 = \frac{V_{in}(C_{Q3} + C_{Q4})}{nI} \quad (3)$$

where  $I$  is the load current;  $C_{Q3}$  and  $C_{Q4}$  are the capacitance of the two intrinsic capacitors of Q3 and Q4, respectively;  $L_r$  is the resonant inductance.  $I_c$  can be calculated as follows:

$$I_c = V_{in} \sqrt{\frac{(C_{Q1} + C_{Q2})(J^2 - 1)}{L_r}} \quad (4)$$

where

$$J = \frac{nI}{V_{in}} \sqrt{\frac{L_r}{(C_{Q1} + C_{Q2})}} \quad (5)$$

The controller determines the duty cycle based on the input voltage and load current to achieve the desired output voltage. As shown in the above equations, the control law is highly nonlinear and as the result, it leads to a very complicated design in conventional approach.

The controller determines the duty cycle based on the input voltage and load current to achieve the desired output voltage. As shown in the above equations, the control law is highly nonlinear and as the result, it leads to a very complicated design in conventional approach. A Simulink model based on the above equations is developed and shown in Fig. 2.

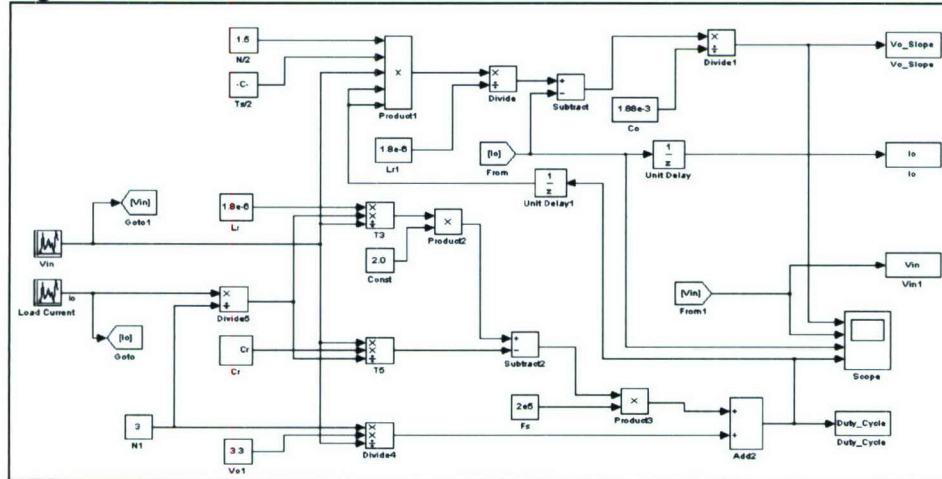


Fig. 2. The Simulink Model

### 3. The Neural Network Controller Design

The Simulink model in Fig. 2 describes the relationship between the input voltage, load current, and duty cycle. It is well-known that artificial neural networks can approximate any nonlinear function to the desired accuracy, after it is fully trained. In this research, the Simulink model is used to obtain data set for neural network training and testing. A feed-forward neural network with one hidden layer is employed to calculate the desired duty cycle. There are six neurons in the hidden layer and one neuron in the output layer. The weights of the neural network ( $W$ ) are initialized at random, and then updated by back-propagation algorithm. To speed up the training, Levenberg-Marquardt algorithm is employed:

$$W(k+1) = W(k) + \Delta W \quad (6)$$

where

$$\Delta W = (J_a^T J_a + \mu I)^{-1} J_a^T e \quad (7)$$

where  $J_a$  is the first order derivative of the error function respect to the neural network weight (also called the Jacobian matrix),  $e$  is the output error (i.e., the difference between the neural network output and the desired output),  $\mu$  is a learning parameter, and  $k$  is the number of iterations.

The training error is reduced to the desired accuracy after about 100 iterations, as shown in Fig. 3.

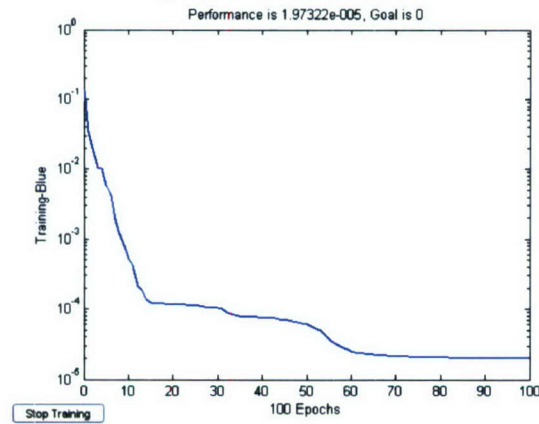


Fig. 3. Neural Network Training Error

Fig. 4 shows the duty cycle estimation of the neural network and the desired value. It is shown that once the neural network is trained, it is able to change the duty cycle based the change of the input voltage.

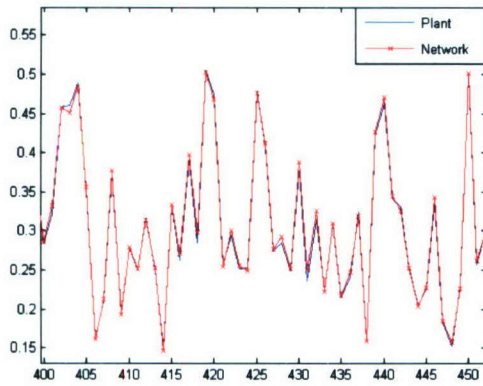


Fig. 4. Duty Cycle Estimation

#### 4. Experimental Results

In order to investigate the performance of the neural network controller from experimental data, a DSP evaluation module eZdsp® F2812 and a phase-shifted zero-voltage-switching evaluation board UCC3895EVM are used to implement the neural controller and the power circuit, respectively. The on-board digital signal processor TMS320F2812 is a 32-bit CPU with 150 MIPS (million instructions per second) operating speed. The input voltage of the phase-shifted full-bridge converter ranges from 36V to 72V, with 48V as its nominal input. The nominal value of the output of the converter is 3.3V, with nominal load current. The neural controller is developed in C language.

First, the analog signals from the power circuit are digitized and processed before being fed into the neural network's inputs. Since the power circuit is driven by 400 kHz high frequency PWM signals, the raw output voltage signal is quite noisy during the switch turn-on and turn-off time. To solve this problem, multiple samples are taken to obtain the average value over a certain period of time. Next, the processed data is fed into the neural network to estimate the desired duty cycle. The new estimated duty cycle is then applied to the power circuit, and the output voltage is monitored.

At load current 2A, the converter's output voltage is measured under different input voltages and the result is shown in Figure 5. The desired nominal output voltage is 3.3V by design. As shown in the plot, the maximum output is 3.36V, which results in only 2% error rate.

Line regulation is a measure of the ability of the power supply to maintain its output voltage, given changes in the input line voltage. Here, we define the percentage of line regulation as:

$$Line\_Reg = \frac{|V_{0(Low\_in)} - V_{0(High\_in)}|}{V_{0(nominal)}} \times 100\%$$

where Line\_Reg represents the percentage of line regulation,  $V_{0(Low\_in)}$  represents the output voltage when the input voltage is low (i.e., 36V), and  $V_{0(High\_in)}$  represents the output voltage when the input voltage is high (i.e., 60V), and  $V_{0(nominal)}$  represents the output voltage when the input voltage is at its nominal value (i.e., 48V). Experimental

results show the percentage of line regulation is 0.9%, which concludes that the neural controller provides excellent line regulation.

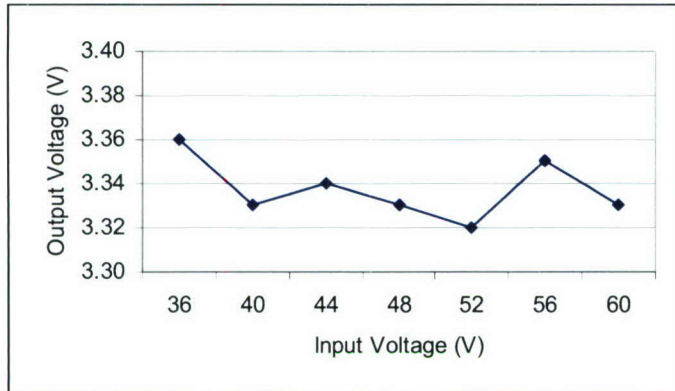


Figure 5. Line Regulation at 2A Load Current

The 10% to 90% load step up and 90% to 10% load step down transient response of the overall system is captured on the oscilloscope as shown in Figure 6. The voltage peak is between 0.4 to 0.5V for both load step up and step down transient response, which is well within the normal range of switching regulator's transient performance.

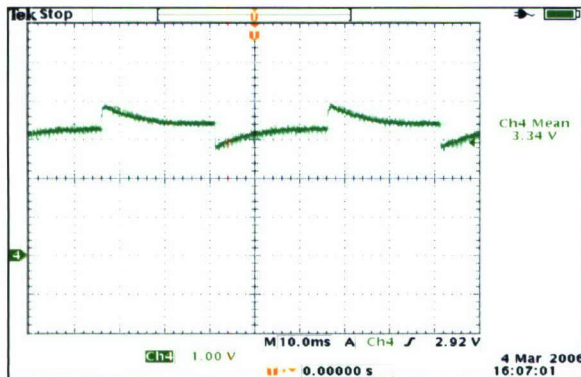


Figure 6. 10%  $\leftrightarrow$  90% Load Transient Response at 60V

## 5. Conclusion

A neural controller as an alternative to classic controller for PSFB DC-DC converter is proposed in this paper. The simulation shows that the neural network controller is able to estimate the desired duty cycle under several dynamic conditions. More research works will be done to implement the neural network controller in hardware.

## References

1. P. R. Chetty, "Resonant power supplies: Their history and status," IEEE Aerosp. Electron. Syst. Mag., vol. 7, no. 4, pp. 23–29, Apr. 1992.
2. M. G. Kim and M. J. Youn, "An energy feedback control of series resonant converters," IEEE Trans. Power Electron., vol. 6, no. 4, pp. 338–345, Jul. 1991.
3. J. M. Carrasco, E. Galván, G. E. Valderrama, R. Ortega, and A. Stankovic, "Analysis and experimentation of nonlinear adaptive controllers for the series resonant converter," IEEE Trans. Power Electron., vol. 15, no. 3, pp. 536–544, May 2000.
4. Hang-Seok Choi and Bo Hyung Cho, "Novel Zero-Current-Switching (ZCS) PWM Switch Cell Minimizing Additional Conduction Loss," IEEE Trans. On Industrial Electronics, Vol. 49, No. 1, P. 165-172, February 2002
5. Hang-Seok Choi, Jung-Won Kim, and Bo Hyung Cho, "Novel Zero-Voltage and Zero-Current-Switching (ZVZCS) Full-Bridge PWM Converter Using Coupled Output Inductor," IEEE Trans. On Power Electronics, Vol. 17, No. 5, P. 641-648, September 2002
6. Xinbo Ruan and Yangguang Yan, "A Novel Zero-Voltage and Zero-Current-Switching PWM Full-Bridge Converter Using Two Diodes in Series With the Lagging Leg," IEEE Trans. On Industrial Electronics, Vol. 48, No. 4, P. 777-785, August 2001
7. José M. Quero, Juan M. Carrasco, and Leopoldo G. Franquelo, "Implementation of a Neural Controller for the Series Resonant Converter," IEEE Trans. On Industrial Electronics, Vol. 49, No. 3, P. 628-639, June 2002
8. Farrukh Kamran, Ronald G. Harley, Bruce Burton, Thomas G. Habetler, and Martin A. Brooke, "A Fast On-Line Neural-Network Training Algorithm for a Rectifier Regulator," IEEE Trans. On Power Electronics, Vol. 13, No. 2, March 1998
9. Byungcho Choi, Jaeyeol Kim, Bo H. Cho, Seungwon Choi, and Carl M. Wildrick, "Designing Control Loop for DC-to-DC Converters Loaded With Unknown AC Dynamics," IEEE Trans. On Industrial Electronics, Vol. 49, No. 4, P. 925-932, August 2002
10. Texas Instrument Inc, "BiCMOS Advanced Phase Shift PWM Controller," Data Sheet of UCC3895, January 2001

## Publications and Presentations Related to Project

### Full publications

Copies of full publications follow.

# A neural network based model for blade vortex interaction noise of rotorcrafts

Johnny Fu and Xiao-Hua Yu  
Dept. of Electrical Engineering  
California Polytechnic State University  
San Luis Obispo, CA 93407, USA  
E-mail: [xhyu@calpoly.edu](mailto:xhyu@calpoly.edu)

**Abstract:** This paper focuses on a neural network based system identification model for rotorcraft acoustic data analysis. The original experimental data is recorded by the microphones mounted inside the wind tunnel at NASA Ames Research Center, Moffett field, CA. Feed-forward neural network is employed and different training algorithms are studied. Satisfactory computer simulation results are obtained.

**Keywords:** Neural networks, rotorcraft acoustic data, system modeling

## 1. Introduction

In the studies of rotorcraft flight control and acoustics, the blade vortex interaction (BVI) noise level is an important measure of system performance. This level represents the acoustic noise generated from rotor interaction with the aerodynamic forces generated by previous rotor blade passages, and is the noise that is of the greatest concern for both military and civilian use [1].

BVI noise is closely related with some flight parameters, such as the rotor thrust, the angle of attack, the speed, etc. If pilots are given enough information to avoid certain flight conditions under which a rotorcraft generates the greatest amount of noise, rotorcraft acoustic noise could be dramatically reduced in actual flight. This in turn may help rotorcrafts gain greater acceptance in use by the general public. However, it is shown that BVI noise is a complicated nonlinear function of flight parameters; and thus it is difficult to find a closed-form solution to the problem.

It is well known that artificial neural networks can be used to perform tasks such as machine learning, association, classification, generalization and optimization [2]. An artificial neural network (ANN) is composed of many non-linear adaptive processing elements and is capable of approximating any function under certain conditions. ANN has been widely used in many fields such as system identification, adaptive control, and statistical modeling.

There have been several developments in the past several years in the application of neural networks for wind tunnel experiments, many of which have been initiated by NASA Ames Research Center, Moffett field, CA [4] [5] [6].

In this research, feedforward neural networks are employed to model the acoustic data gathered from two wind tunnel tests for two different types of rotorcrafts (a full scale XV-15 tilt-rotor and a modern 4-bladed rotor). For each type of rotorcraft, a high BVI noise case and a low BVI noise case are considered. Different training methods are applied and compared, including the back-propagation and the extended Kalman filter algorithm. Computer simulations are performed and satisfactory results are obtained.

This paper is organized as follows. Section 2 gives a brief description on the data acquisition procedures of wind tunnel tests. In section 3, the neural network is applied to model the rotorcraft acoustic data, with comparison on different training algorithms. Section 4 concludes the paper and discusses future works for this study.

## 2. The wind tunnel tests

The experimental data used in this study comes from two wind tunnel tests conducted in the 80 x 120-foot wind tunnel at NASA Ames Research Center at Moffett Field, CA. The purpose of this test was to determine the relationship between the acoustic noise level and different flight conditions (Fig. 1).

The acoustic data is acquired in the wind tunnel by microphones mounted in strategic locations that are estimated to be the locations of maximum noise relative to the location of the test rotor. Several of these microphones are also mounted on a traverse that allows it to be moved horizontally in the wind tunnel. Individual data points are characterized by a set of parameters, including the rotor thrust  $C_T/\sigma$ , a measure of the rotor's lift capability; the advance ratio  $\mu$ , which is

the ratio of true forward airspeed over the wind speed at rotor tip; the rotor tip-path-plane angle  $\alpha_{TPP}$ , also commonly known as the angle of attack; and finally, the distance (measured in inches) between the microphone and test rotor  $x_m$ . For one set of flight parameters, the acoustic data is sampled and recorded for 16 revolutions of the rotor, at the sampling frequency of 2048 samples per second. The entire data is then averaged to obtain one set of data for a single revolution. The BVI data is measured in Pascals. Figure 2 shows an example of the placement of microphones relative to a test rotor in a wind tunnel experimental setup.

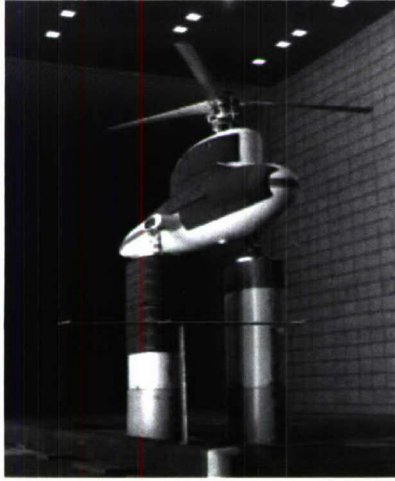


Figure 1: The wind tunnel test setup

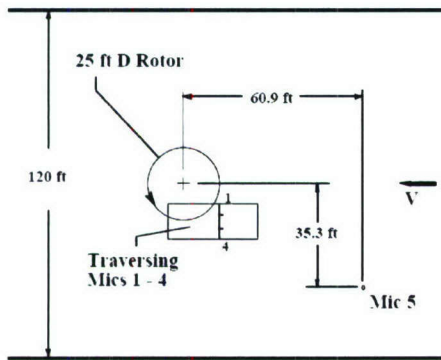


Figure 2: Placement of microphones in wind tunnel experiment

### 3. Neural network based model for acoustic noise level measurement

A block diagram for neural network based system identification is shown in Fig. 3. A training data set which contains input and desired output (target) is provided; then the weights of the neural

network are adjusted to minimize a performance index until it is less than a pre-set threshold.

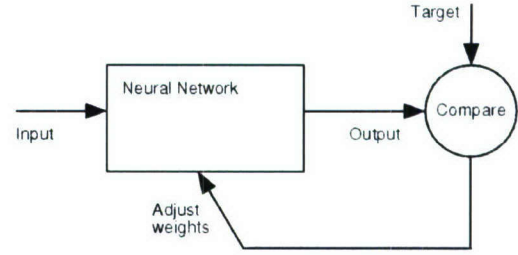


Figure 3: The neural network based system identification

Mean-square-error (MSE) is the most commonly used criterion and is defined as:

$$J = \frac{1}{NK} \left[ \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K (z_{kn} - d_{kn})^2 \right] \quad (1)$$

where  $N$  is the number of input/output pairs in the training set and  $K$  is the number of outputs of the neural network;  $z_{kn}$  is the  $k^{\text{th}}$  network output for the  $n^{\text{th}}$  input/output pair in the training set;  $d_{kn}$  represents the  $k^{\text{th}}$  desired (target) output for the  $n^{\text{th}}$  input/output pair. When this performance function is at a minimum, the output of a neural network fits the training data as closely as possible.

In this research, three different training algorithms are considered. The first one is a multi-layer ADALINE (adaptive linear neuron) neural network trained by LMS (least-mean-square) algorithm. The neural network contains 5 neurons in the input layer for flight parameters and 1 output for the BVI noise. All the neurons are linear and there is no hidden layer in this case. For the other two neural networks, each of them is constructed with a hidden layer of 10 nonlinear neurons (sigmoid activation function) and a linear output neuron.

The weight of neural network can be adjusted using:

$$w_m = w_{m-1} + \Delta w_m \quad (2)$$

where  $m$  is the number of training iteration.

Learning in back-propagation is performed in two steps, i.e., the forward pass and the backward pass. In the latter, the error signal (i.e., the difference between the neural network output and the target) is feedback from the output to the input, layer by layer, and all the weights are adjusted in proportion to this error, i.e.,

$$\Delta w_m = \eta \frac{\partial J}{\partial w} \quad (3)$$

Where  $\eta$  is called the learning/training rate;  $J$  is the objective function defined in Eq. (1) and  $W$  is

the weight matrix of the neural network. For sigmoid function, we have:

$$g(u) = \frac{1}{1 + e^{-u}} \quad (4)$$

where  $g$  is the output of the nonlinear neuron and

$$u = WX$$

with weight matrix  $W$  and input to the neuron  $X$  (both vectors). Therefore,

$$\frac{\partial J}{\partial W} = (z - d)u(1 - u)X \quad (5)$$

For the sake of simplicity, the index of neuron is not included in the above formula.

It is well known that the extended Kalman filter algorithm can be used to estimate the state of a nonlinear system. If the neural network is considered as a nonlinear system, then the weights can be updated through the Kalman gain  $K$  in each training iteration  $m$ :

$$w_m = w_{m-1} + K_m [z - d] \quad (6)$$

$$K_{m+1} = P_m H_m^T [R + H_m^T P_m H_m]^{-1} \quad (7)$$

$$P_{m+1} = P_m - K_m H_m^T P_m \quad (8)$$

where  $R$  is the covariance of the identification error. Again for the sake of simplicity, the index of neuron is not included in the above formula.

### 3.1. The neural network based model for XV-15

The wind tunnel test data was recorded at the following conditions for the low BVI case of the XV-15 tilt-rotorcraft:  $C_T/\sigma = 0.0756$ ,  $\mu = 0.125$ ,  $\alpha_{TPP} = -15.04^\circ$ . Three neural networks with different training algorithms are used here for simulation (i.e., the ADALINE network with a learning rate of 0.1, the back-propagation trained neural network with a learning rate of 0.1, and the Kalman filter trained neural network an  $S$  value (error covariance) of  $8e-5$ ). The simulation results are presented in Fig. 4. It is shown that the Kalman filter trained neural network yields the best performance (approximately 5 times better than the back-propagation algorithm in terms of mean square error).

The high BVI case of the XV-15 tilt-rotorcraft is recorded under the following conditions:  $C_T/\sigma = 0.0913$ ,  $\mu = 0.170$ ,  $\alpha_{TPP} = -3.08^\circ$ , and  $x_m = 75$  inches. Using the same neural network that employed for the low BVI case, we obtained comparable results, as shown in Fig. 5.

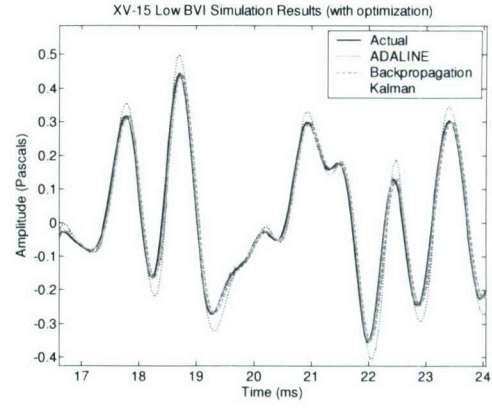


Figure 4: Simulation results on XV-15 rotorcraft (low-BVI case)

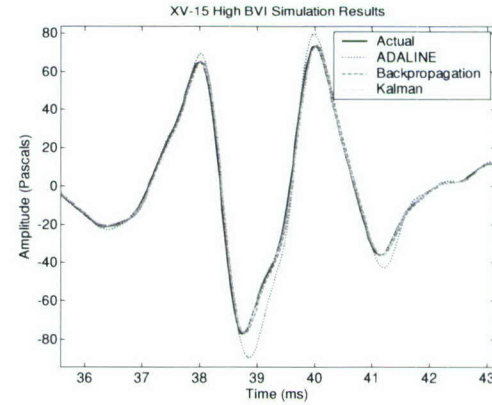


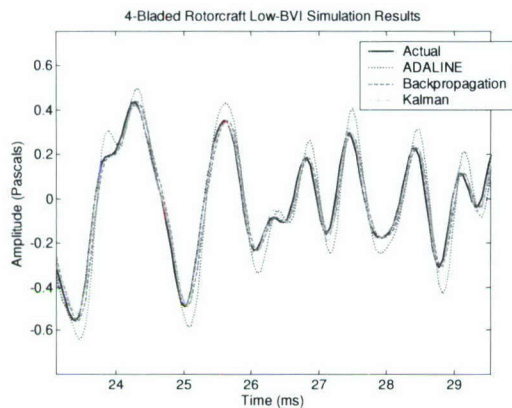
Figure 5: Simulation results on XV-15 rotorcraft (high-BVI case)

### 3.2. The neural network based model for a modern 4-bladed rotorcraft

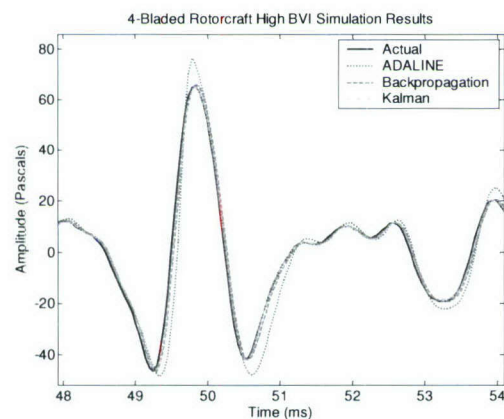
For the modern 4-bladed rotorcraft, the acoustic level data was recorded at the following wind tunnel conditions:  $C_T/\sigma = 0.04$ ,  $\mu = 0.1$ ,  $\alpha_{TPP} = -10^\circ$ , and  $x_m = 364$  inches. Again, three neural networks with different training algorithms are used here for simulation (i.e., the ADALINE network with learning rate set at 0.5, the back-propagation trained neural network with 10 hidden neurons and a learning rate of 0.5, and the Kalman filter trained neural network with 10 hidden neurons and an  $S$  value (error covariance) of  $5e-4$ ). The simulation results are shown in Fig. 6. It is shown that the extended Kalman filter training algorithm outperforms the other two networks, with the minimum amount time delay and the lowest MSE (mean-square-error).

The final case involves the high BVI case of the modern 4-bladed rotorcraft. The measurement data is taken under the following test conditions:  $C_T/\sigma = 0.09$ ,  $\mu = 0.19$ ,  $\alpha_{TPP} = 4^\circ$ , and  $x_m = -70$  inches. From the simulation result shown in Fig. 7,

we find that the extended Kalman filter training algorithm again yields the best overall performance with MSE being only 25% of the MSE for the back-propagation algorithm. The ADALINE algorithm has the worst performance and produces larger errors on the BVI peaks.



**Figure 6: Simulation results on 4-bladed rotorcraft (low-BVI case)**



**Figure 7: Simulation results on 4-bladed rotorcraft (high-BVI case)**

## 4. Conclusion

The blade vortex interaction (BVI) noise level is an important measure in the studies of rotorcraft acoustics. Successful modeling and prediction of BVI noise levels may allow for the development of future flight systems that actively control the emission of BVI noise or alert rotorcraft pilots when BVI noise levels are excessively high in noise-sensitive areas or trajectories.

In this research, artificial neural networks are successfully applied to model BVI noise levels and compared with the actual data gathered from wind tunnel data. The simulation results showed that a trained neural network is able to capture the

relationship between the different flight parameters and the corresponding BVI noise level. Future works that may evolve from this study include using neural networks to analyze other sets of rotorcraft data for the purposes of modeling and prediction.

## Acknowledgements

The author thanks Dr. William Warmbrodt, Cahit Kitaplioglu, and the other members of the Aeromechanics Branch at NASA Ames Research Center for supporting this work. This work was also partially sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-04-1-0436.

## References

- [1] Johnson, W., *Helicopter Theory* (Princeton, NJ: Princeton University Press, 1980).
- [2] Zaknich, A., *Neural Network for Intelligent Signal Processing* (Singapore: World Scientific, 2003).
- [3] Schalkoff, R. J., *Artificial Neural Networks* (New York: McGraw-Hill, 1997).
- [4] Kitaplioglu, C., *Blade-Vortex Interaction Noise of a Full-Scale XV-15 Rotor Tested in the NASA Ames 80- by 120-Foot Wind Tunnel*, NASA Technical Memorandum 208789, Moffett Field: National Aeronautics and Space Administration.
- [5] Lo, and Schmitz, F., *Model-Based Neural Networks for Rotorcraft Ground Noise Prediction*, Proceedings of the 38<sup>th</sup> AIAA Aerospace Sciences Meeting & Exhibit, Reno, NV, January 10-13, 2000.
- [6] Holger, G., et al. *Neural Networks For BVI System Identification*, Proceedings of the 29<sup>th</sup> European Rotorcraft Forum, Friedrichshafen, Germany, September 16-18, 2003.
- [7] Boris, I., and Hoaglin, D., *How to Detect and Handle Outliers* (Milwaukee, Wisc: ASQC Quality Press, 1993).
- [8] Fu, J., *Modeling and analysis of rotorcraft acoustic data with neural networks*, M. S. thesis, June 2004.

# Rotorcraft Acoustic Noise Estimation and Outlier Detection

Johnny Fu, and Xiao-Hua Yu, *Member, IEEE*

**Abstract** -- This paper focuses on the application of artificial neural networks for rotorcraft acoustic data modeling, prediction, and outlier detection. The original data is recorded by microphones mounted inside a wind tunnel at NASA Ames Research Center, Moffett Field, CA. The experimental data is first acquired in the time-domain as a time history measurement; then the sound pressure level (SPL) that represents the acoustic noise in frequency domain is derived from the time history dataset. In this study, neural networks based models are developed in both time domain and frequency domain. Outlier detection is then performed using modified Z-scores for SPL data to find test points that are statistically inconsistent with the neural network model. Satisfactory computer simulation results are obtained.

## I. INTRODUCTION

Wind tunnel acoustic tests of rotorcraft can be conducted to determine the conditions under which a rotorcraft generates the greatest amount of noise. If pilots are given enough information to avoid these flight conditions, rotorcraft acoustic noise could be reduced in actual flight. This in turn may help rotorcrafts gain greater acceptance in use by the general public. The specific type of noise that rotorcraft researchers are generally interested in for noise reduction is the blade-vortex interaction noise (BVI). This is the noise that occurs when a rotorcraft blade passes through the aerodynamic wake of previous blades, and is the noise that is of the greatest concern for both military and civilian use [1]. Successful modeling and prediction of BVI noise levels may allow for the development of future flight systems that actively control the emission of BVI noise or alert rotorcraft pilots when BVI noise levels are excessively high in noise-sensitive areas or trajectories. However, due to the highly nonlinear relationship between flight conditions and BVI noise level, the closed-form solution is very difficult to obtain with aerodynamics and acoustic analysis.

It is well known that artificial neural networks can be used to perform tasks such as machine learning, association, classification, generalization and optimization [2] [3]. An artificial neural network (ANN) is composed of many nonlinear adaptive processing elements and is capable of approximating any function under certain conditions. ANN has been widely used in the field of system identification, adaptive control, and statistical modeling. There have been some developments in the past several years in the

application of neural networks for wind tunnel experiments, such as [4], [5], and [6]. Kottapalli and Kitaplioglu investigated BVI data quality analysis and prediction for the XV-15 tilt-rotorcraft using backpropagation and radial basis function neural networks [4]. Lo and Schimtz demonstrated the prediction of BVI sound pressure levels using neural networks and the Higher Harmonic Control Aeroacoustics Rotor Test (HART) database [5]. Two parameters that are related with flight conditions, including rotor tip-path-plane angle and advance ratio, are considered. Holger et al. used neural networks to model data time histories gathered from in-flight tests conducted by Eurocopter Deutschland on the BO 105 helicopter [6]. The BVI noise is modeled as a function of phase angles sent to the helicopter's individual blade controller.

In this research, feedforward neural networks will be employed to model the acoustic data gathered from a wind tunnel test of a 3-bladed rotorcraft at NASA Ames Research Center. The relationship between the BVI noise and multiple flight conditions are considered here, including the rotor thrust, the advance ratio, the angle of attack of rotor, and the distance between the microphone and test rotor. Two typical situations are considered here, i.e., a high BVI noise case and a low BVI noise case. Different training methods are applied and compared, including backpropagation and the extended Kalman filter algorithm. The trained neural networks in this study is then be used to predict BVI noise levels not included in the training dataset. This prediction test will demonstrate a trained neural network's ability to estimate/interpolate data that was not tested in the wind tunnel. To detect outliers in the wind tunnel data, the median-based modified Z-score method is employed. An outlier is defined as the data point that is statistically inconsistent with the expectations about this point based on the model derived using the data set [7]. Points that deviate greatly from the model could indicate a fault in the hardware data acquisition system or present new information about the BVI noise level data not yet incorporated into the neural network model. Comparing with the conventional method which is based on the average value of the dataset and may be heavily influenced by the presence of outliers, the modified Z-score method is more robust and thus more appropriate for outlier detection.

This paper is organized as follows. First, a brief description on the wind tunnel test setup and data collection is given in section 2. In Section 3, the neural network based model in time-domain is developed, with comparison on different training algorithms. Section 4 concentrates on the model of BVI noise level in frequency-domain. The results

Johnny Fu is with Sierra Lobo, Inc., Moffett Field, CA, USA (e-mail: jfu@mail.arc.nasa.gov).

Xiao-Hua Yu is with the Department of Electrical Engineering, California Polytechnic State University, San Luis Obispo, CA 93407, USA (e-mail: xhyu@ee.calpoly.edu).

from a study to identify outliers in the dataset are presented in Section 5. Finally, a conclusion is presented in Section 6.

## II. DATA COLLECTION AND TEST SETUP

The experimental data used in this study comes from a wind tunnel test of a 3-bladed rotorcraft conducted in the 80 x 120-foot wind tunnel at NASA Ames Research Center at Moffett Field, CA (Fig. 1). The purpose of this test was to determine the relationship between the acoustic noise level and different flight conditions.

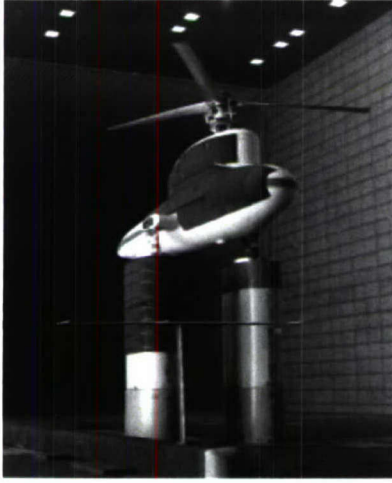


Figure 1. The wind tunnel test setup

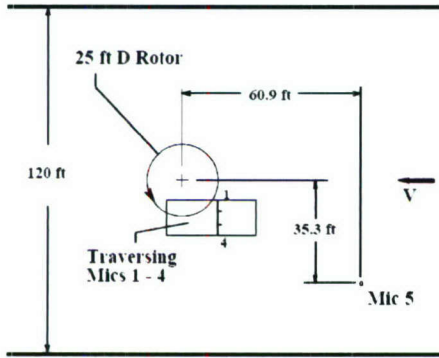


Figure 2. Placement of microphones in wind tunnel experiment

The acoustic data is acquired in the wind tunnel by microphones mounted in strategic locations that are estimated to be the locations of maximum noise relative to the location of the test rotor. Several of these microphones are also mounted on a traverse that allows it to be moved horizontally in the wind tunnel. Individual data points are characterized by a set of parameters, including the rotor thrust  $C_T/\sigma$ , a measure of the rotor's lift capability; the advance ratio  $\mu$ , which is the ratio of true forward airspeed over the wind speed at rotor tip; the rotor tip-path-plane

angle  $\alpha_{TPP}$ , also commonly known as the angle of attack; and finally, the distance (measured in inches) between the microphone and test rotor  $x_m$  [4]. Figure 2 shows an example of the placement of microphones relative to a test rotor in a wind tunnel experimental setup.

## III. NEURAL NETWORK BASED MODEL IN TIME-DOMAIN

A block diagram for neural network based system identification is shown in Fig. 3. A training data set which contains input and desired output (target) is provided; then the weights of the neural network are adjusted to minimize a performance index until it is less than a pre-set threshold.

Mean-square-error (MSE) is the most commonly used criterion and is defined as:

$$J = \frac{1}{NK} \left[ \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K (z_{kn} - d_{kn})^2 \right] \quad (1)$$

where  $N$  is the number of input/output pairs in the training set and  $K$  is the number of outputs of the neural network;  $z_{kn}$  is the  $k^{\text{th}}$  network output for the  $n^{\text{th}}$  input/output pair in the training set;  $d_{kn}$  represents the  $k^{\text{th}}$  desired (target) output for the  $n^{\text{th}}$  input/output pair. When this performance function is at a minimum, the output of a neural network fits the training data as closely as possible.

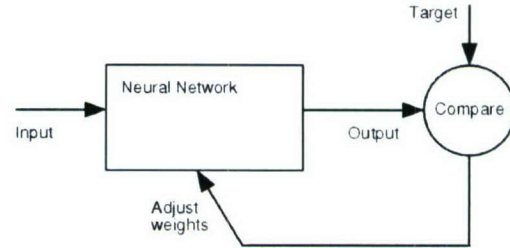


Figure 3. The neural network based system identification

In this research, three different training algorithms are considered. The first one is a multi-layer ADALINE (adaptive linear neuron) neural network trained by LMS (least-mean-square) algorithm. The neural network contains three neurons in the input layer for previous time history values and one output for the BVI noise. All the neurons are linear and there is no hidden layer for this neural network. For the other two neural networks, trained by back-propagation and Kalman filter algorithm respectively, each of them is constructed with a hidden layer of 10 nonlinear neurons (sigmoid activation function) and a linear output neuron.

The weight of neural network can be adjusted using:

$$w_m = w_{m-1} + \Delta w_m \quad (2)$$

where  $m$  is the number of training iteration.

Learning in back-propagation is performed in two steps, i.e., the forward pass and the backward pass. In the latter, the error signal (i.e., the difference between the neural network output and the target) is feedback from the output to the input, layer by layer, and all the weights are adjusted in proportion to this error, i.e.,

$$\Delta w_m = \eta \frac{\partial J}{\partial W} \quad (3)$$

where  $\eta$  is called the learning/training rate;  $J$  is the objective function defined in Eq. (1) and  $W$  is the weight matrix of the neural network. For sigmoid function, we have:

$$g(u) = \frac{1}{1 + e^{-u}} \quad (4)$$

where  $g$  is the output of the nonlinear neuron and

$$u = WX$$

with weight matrix  $W$  and input to the neuron  $X$  (both vectors). Therefore,

$$\frac{\partial J}{\partial W} = (z - d) u(1 - u) X \quad (5)$$

For the sake of simplicity, the index of neuron is not included in the above formula.

It is well known that the extended Kalman filter algorithm can be used to estimate the state of a nonlinear system. If the neural network is considered as a nonlinear system, then the weights can be updated through the Kalman gain  $K$  in each training iteration  $m$ :

$$w_m = w_{m-1} + K_m [z - d] \quad (6)$$

$$K_{m+1} = P_m H_m^T [R + H_m^T P_m H_m]^{-1} \quad (7)$$

$$P_{m+1} = P_m - K_m H_m^T P_m \quad (8)$$

where  $R$  is the covariance of the identification error. Again for the sake of simplicity, the index of neuron is not included in the above formula.

In the wind tunnel test, the acoustic level data was recorded at the following wind tunnel conditions for a 3-bladed rotorcraft:  $C_T/\sigma = 0.04$ ,  $\mu = 0.1$ ,  $\alpha_{TPP} = -10^\circ$ , and  $x_m = 364$  inches. Three neural networks with different training algorithms are used here for simulation (i.e., the ADALINE network with learning rate set at 0.5, the back-propagation trained neural network with 10 hidden neurons and a learning rate of 0.5, and the Kalman filter trained neural network with 10 hidden neurons and an  $S$  value (error covariance) of  $5 \times 10^{-4}$ ). The simulation results are shown in Fig. 4. It is shown that the extended Kalman filter training algorithm outperforms the other two networks, with the minimum amount time delay and the lowest MSE (mean-square-error).

The high BVI case of the 3-bladed rotorcraft is also studied in this research. The measurement data is taken under the following test conditions:  $C_T/\sigma = 0.09$ ,  $\mu = 0.19$ ,  $\alpha_{TPP} = 4^\circ$ , and  $x_m = -70$  inches. From the simulation result shown in Fig. 5, we find that the extended Kalman filter

training algorithm again yields the best overall performance with MSE being only 25% of the MSE for the back-propagation algorithm. The ADALINE algorithm has the worst performance and produces larger errors on the BVI peaks.

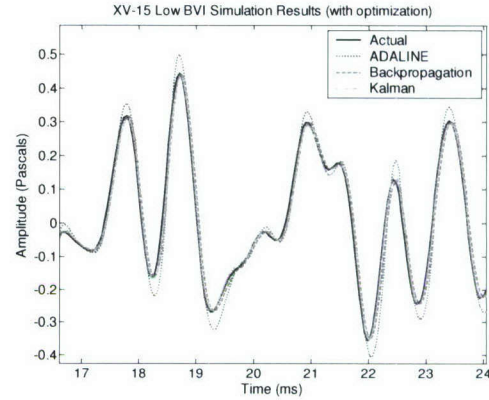


Figure 4. Simulation results on 3-bladed rotorcraft (low-BVI case)

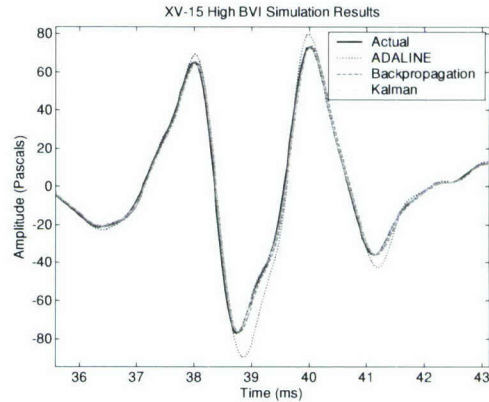


Figure 5. Simulation results on 3-bladed rotorcraft (high-BVI case)

The simulation results with different training algorithms are summarized in table 1, where the percentages of errors are shown for each situation:

TABLE I  
SIMULATION RESULTS TABLE FOR DIFFERENT TRAINING ALGORITHMS

Algorithm	Low BVI	High BVI
ADALINE	14.25 %	13.45 %
Backpropagation	7.45 %	7.04 %
Kalman	2.1 %	2.95 %

#### IV. BVI NOISE LEVEL ESTIMATION

The sound pressure level metric provides a convenient way for rotorcraft researchers to study the overall BVI

acoustic noise for a given set of test parameters. To obtain a single-value representing the overall BVI noise level from one rotor revolution, the time-domain data is transformed into the frequency domain using a FFT. The resulting power spectrum is integrated and converted from units of Pascals<sup>2</sup> to units of dB to obtain a BVI noise level value that characterizes the acoustic noise from one revolution of the rotor. A digital bandpass filter is applied to highlight the BVI frequencies.

The whole dataset contains about 1000 total BVI sound pressure level measurements for 4 microphones. The database was divided into three approximately equal datasets for training, validation, and testing. The inputs to the neural network include  $C_T/\sigma$ ,  $\mu$ ,  $\alpha_{TPP}$ , the microphone number, and the traverse position (in inches). The neural network contains 10 hidden neurons, and is trained off-line. To speed up the training, the extended-bar-delta-bar adaptive learning rate algorithm is used:

$$\Delta w_m = \rho \Delta w_{m-1} - (1 - \rho) \alpha_m \delta_m \quad (9)$$

and  $\delta_m = \partial J / \partial W$

where  $\rho$  is the momentum factor (commonly 0.9 for most applications) and  $\alpha_m$  is called the adaptive learning rate. A mean square error of  $6 \times 10^{-4}$  is achieved in approximately 10,000 training epochs.

In the studies of BVI sound pressure level, a contour plot which shows the relationship between BVI and the variations of two flight parameters is very important. However, in the case when there are not enough points in the dataset, the contour plot may not display any recognizable pattern due to the lack of measurement. Fig. 6 shows one example when the two flight parameters  $\alpha_{TPP}$  (the rotor tip-path-plane angle, or the angle of attack) and the advance ratio  $\mu$  at the rotor tip.

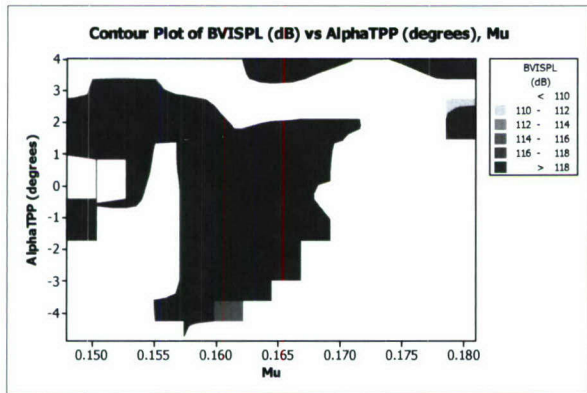


Figure 6. Relationship between the BVI noise level vs.  $\mu$  and  $\alpha_{TPP}$

Using the trained neural network, it is possible to interpolate the incomplete dataset with the points predicted by the neural network. A complete dataset can then be

constructed and a contour plot which is more visible can be generated (Fig. 7).

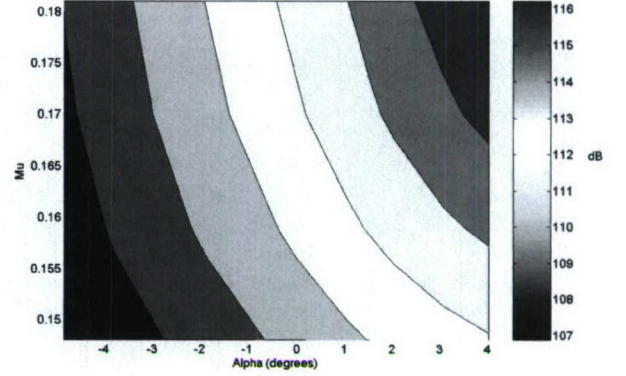


Figure 7. Relationship between the BVI noise level vs.  $\mu$  and  $\alpha_{TPP}$  (reconstructed with neural network model)

## V. OUTLIER DETECTION

The conventional method for outlier detection is based on the average value of the dataset and may be heavily influenced by the presence of outliers. In this research, the robust modified Z-score method is used to find points where there the residual between the estimated neural network output and the recorded BVI noise level are outside the range of residuals for the bulk of the data. To determine this range of residuals, a benchmark median absolute deviation from the median residual value is calculated (M.A.D.) using the prediction residuals of the trained neural network. The equation for this M.A.D. value is shown in Eq. (10) where the residuals are contained in a set  $\{e\}$ ,  $e_i$  is the residual for the  $i^{\text{th}}$  data point, and  $\bar{e}$  is the median residual:

$$M.A.D. = \text{median}\{|e_i - \bar{e}|\} \quad (10)$$

Once the benchmark M.A.D. is calculated, a modified Z-score ( $M$ ) is calculated for each prediction residual using Eq. (11). Here, the scaling factor 0.6745 is used to normalize the expected value of the M.A.D. so that it is equal to the standard deviation in a normal distribution.

$$M_i = \frac{0.6745 (e_i - \bar{e})}{MAD} \quad (11)$$

The absolute value of the modified Z-score for each point is compared to the M.A.D. to determine whether it is an outlier point or not. When  $|M_i| > \psi \cdot (M.A.D.)$ , the  $i^{\text{th}}$  data point is identified as an outlier. When  $\psi = 3.5$  in normally distributed data, approximately 5% of the data points are identified as outliers.

For the BVI noise level data, an analysis of the outliers in the original 309 point testing dataset is conducted. The MAD of the validation residuals was calculated to be 1.57 dB. Using an outlier factor of 3.5, 6 outliers were detected

in the data. The residuals from these outlier points ranged from -5.73 dB to 6.67 dB. Three of these points came from a common set of input parameters ( $C_T/\sigma = 0.0754$ ,  $\mu = 0.125$ , and  $x_m = 225$ ). Such a finding would indicate the need for closer inspection of the wind tunnel test log for these points to see if any unusual occurrences were recorded for these data points. Lowering the outlier factor to 2.5 increases the number of outliers detected to 18 points. These points have absolute residuals  $> 4$  dB. It is noticed that several of these outliers were different microphones with the same set of wind tunnel operating parameters. This means that for these data points, the neural network predicted values and the wind tunnel values compared poorly and thus the data for those points require a closer inspection.

## VI. CONCLUSION

In this research, artificial neural networks are successfully employed to estimate BVI noise levels and compared with the actual data gathered from wind tunnel data in both time and frequency domain. The simulation results showed that a trained neural network is able to capture the relationship between the different flight parameters and the corresponding BVI noise level in frequency domain. Outlier detection was performed using modified Z-scores and a benchmark M.A.D. statistic calculated from the neural network prediction residuals. Only a few outliers were detected in the non-training portion of the dataset, indicating that the overall wind tunnel data was consistently recorded.

Future works that may evolve from this study include using the neural network to analyze other sets of rotorcraft data. Other data sets may be analyzed following the methodology of this study for the purposes of modeling, prediction, and outlier detection.

## ACKNOWLEDGEMENTS

The author would like to thank Dr. William Warmbrodt, Cahit Kitaplioglu, and the other members of the Aeromechanics Branch at NASA Ames Research Center for supporting this work. This work was also partially sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855.

## REFERENCES

- [1] W. Johnson, *Helicopter Theory* (Princeton, NJ: Princeton University Press, 1980).
- [2] A. Zaknich, *Neural Network for Intelligent Signal Processing* (Singapore: World Scientific, 2003).
- [3] R.J. Schalkoff, *Artificial Neural Networks* (New York: McGraw-Hill, 1997).
- [4] C. Kitaplioglu, Blade-Vortex Interaction Noise of a Full-Scale XV-15 Rotor Tested in the NASA Ames 80- by 120-Foot Wind Tunnel, NASA Technical Memorandum 208789. Moffett Field: National Aeronautics and Space Administration.

- [5] Lo, and Schmitz, F., *Model-Based Neural Networks for Rotorcraft Ground Noise Prediction*, Proceedings of the 38<sup>th</sup> AIAA Aerospace Sciences Meeting & Exhibit, Reno, NV, January 10-13, 2000.
- [6] Holger, G., et al. *Neural Networks For BVI System Identification*, Proceedings of the 29<sup>th</sup> European Rotorcraft Forum, Friedrichshafen, Germany, September 16-18, 2003.
- [7] B. Iglewicz and D.C. Hoaglin, *How to Detect and Handle Outliers* (Milwaukee, WI: ASQC Quality Press, 1993).
- [8] Fu, J., Modeling and analysis of rotorcraft acoustic data with neural networks, M. S. thesis, June 2004.

# A Neural Network Controller for a Class of Phase-Shifted Full-Bridge DC-DC Converter

Weiming Li, Xiao-Hua Yu, and Taufik

Department of Electrical Engineering  
California Polytechnic State University  
San Luis Obispo, CA 93407, USA

xhyu@ee.calpoly.edu

**Abstract** - DC-DC converters can be found in almost every power electronics device. To improve the controller response of a DC-DC converter to dynamical system changes, neural network has been chosen as an alternative to classic methods. However, no prior work has been done in the neural network approach for control of a PSFB (Phase-Shifted Full-Bridge) converter yet. In this research, a multi-layer feed-forward neural network controller is proposed. The neural network based controller has the advantage of adaptive learning ability, and can work under the situation when the input voltage and load current fluctuate. The Levenberg-Marquardt back-propagation training algorithm is used in computer simulation and satisfactory results are obtained.

**Index Terms** - Neural network controller, DC-DC converter

## I. INTRODUCTION

DC-DC converters can be found in almost every electronic device nowadays, since all semiconductor components are powered by DC source. One of the design targets for electronic engineers is to improve the efficiency of power conversion. For PWM (pulse-width modulation) converters, switching loss is an important performance measure. Many different kinds of topologies ([1], [2], [3]) have been investigated in the past to reduce the switching loss. However, those topologies either need additional components for the power circuit, which may introduce some unstable factors to the circuit; or operate at variable frequency, which makes the filter design at output stage very difficult. Phase-shifted zero-voltage switching full-bridge converters overcome the above problems and thus have been received more and more attention recently. It employs zero-voltage-switching (ZVS) technique and allows the voltages across the transistors to swing to zero just before the start of the next conduction cycle ([4], [5], [6]).

Normally, the supply voltage and load current have a wide range of variation; so the controller has to be designed to work under such conditions. The

conventional approach is to assume that the converter is operated around its equilibrium state; then a set of linear equations are derived based on this assumption [9] [10].

Artificial neural networks (ANN) have been widely used in the field of system identification, adaptive control, and statistical modelling in recent years. A neural network is composed of many non-linear adaptive processing elements and is capable of approximating any measurable function under certain conditions.

To improve the controller response of a DC-DC converter to dynamical system changes, neural network has been chosen as an alternative to classic methods ([7], [8]). However, no prior work has been done in the neural network approach for control of a PSFB (Phase-Shifted Full-Bridge) converter yet. In this research, a multi-layer feed-forward neural network controller is proposed. A Matlab Simulink model is developed first to generate the data set; then the neural network is trained by Levenberg-Marquardt back-propagation algorithm. Satisfactory simulation results are obtained.

## II. CIRCUIT ANALYSIS

The circuit diagram of a PSFB DC-DC converter is shown in Fig. 1.

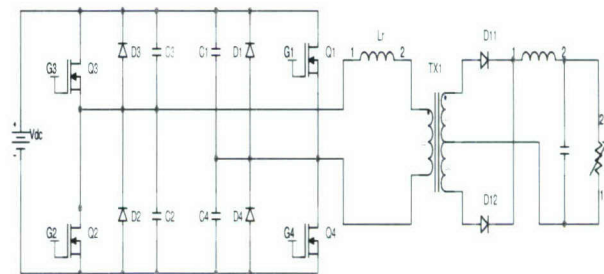


Fig 1. PSFB DC-DC Converter

Based on circuit analysis, the control scheme to drive the switching MOSFET (Q1, Q2, Q3 and Q4) is

very complicated [6]. The circuit is operated in one of the following modes:

Mode 1: The diagonal MOSFET Q3 & Q4 are conducting and power is delivered through the transformer to the load. The primary load current is flowing through the leakage inductance of the transformer. The total primary current is equal to the load current plus increasing magnetizing current of the transformer.

Mode 2: With Q4 on, the capacitance across Q1 is charged to +V. When Q4 turns OFF the current through the transformer inductance starts to charge the drain source capacitance of Q4, while at the same time discharges the capacitance of Q1. This action continues until the body diode of Q1 turns ON to clamp the voltage across Q1 at approximately 0.7V. The current through the transformer is sustained in the upper half of the power circuit.

Mode 3: When the voltage across Q1 reaches approximately 0V, Q1 turns ON. The time that is required for the capacitance of Q4 and Q1 to reach the desired voltage is determined by the characteristics of hardware. The current in this mode circulates through the conduction channels of Q3 and Q1.

Mode 4: Q3 turns off. The transformer current now starts to charge and discharge the capacitance of Q3 and Q2 respectively. It again requires a finite amount of time for the drain voltage of Q2 to reach "0 volt" at which point Q2 is allowed to be switched on. That means non-dissipate turn ON switching is accomplished.

Mode 5: With the complete discharge of its drain source capacitance Q2 is now ready to turn ON. Power is delivered to the load through the conduction path of Q1 and Q2 for an amount of time that is determined by the control circuit. Twice the product of this time, times the operating frequency of the oscillator gives the duty cycle of the converter as in a regular PWM converter. Duty Cycle  $D = 2 \cdot T_{on} \cdot F_s$ .

Mode 6: Following the power transfer by the above diagonal pair, Q1 turns OFF. The voltage across Q4 starts to decrease, and when this voltage reaches 0 volts the next mode starts.

Mode 7: In this mode, Q4 turns on and primary current circulates in the conduction channels of the lower pair.

Mode 8: Q2 turns OFF and current starts to charge and discharge the capacitance of Q2 and Q3 respectively. When the voltage across Q3 has reached 0 volts then Q3 turns ON non-dissipative and the complete cycle repeats itself from Mode 1.

The relationship between the input/output voltage and the duty cycle can be described as:

$$V_0 = \left( \frac{2nV_{in}}{T_s} \right) \left( \Phi \cdot \frac{T_s}{2} - T_1 - \frac{T_2}{2} \right) \quad (1)$$

where  $V_0$  is the output voltage,  $V_{in}$  is the input voltage, both in rms value;  $n$  is the transformer ratio,  $T_s$  is the switching period constant, and  $\Phi$  is the duty cycle.  $T_1$  and  $T_2$  can be obtained from the following equations:

$$T_1 = \frac{(nI + I_c)L_r}{V_{in}} \quad (2)$$

$$T_2 = \frac{V_{in}(C_{Q3} + C_{Q4})}{nI} \quad (3)$$

where  $I$  is the load current;  $C_{Q3}$  and  $C_{Q4}$  are the capacitance of the two intrinsic capacitors of Q3 and Q4, respectively;  $L_r$  is the resonant inductance.  $I_c$  can be calculated as follows:

$$I_c = V_{in} \sqrt{\frac{(C_{Q1} + C_{Q2})(J^2 - 1)}{L_r}} \quad (4)$$

where

$$J = \frac{nI}{V_{in}} \sqrt{\frac{L_r}{(C_{Q1} + C_{Q2})}} \quad (5)$$

The controller determines the duty cycle based on the input voltage and load current to achieve the desired output voltage. As shown in the above equations, the control law is highly nonlinear and as the result, it leads to a very complicated design in conventional approach.

### III. LOAD CURRENT ESTIMATION

Based on the above the analysis, the Simulink model for load current estimation is developed and shown in Fig. 2, which shows the relationship between the change in load current and the change in the output voltage.

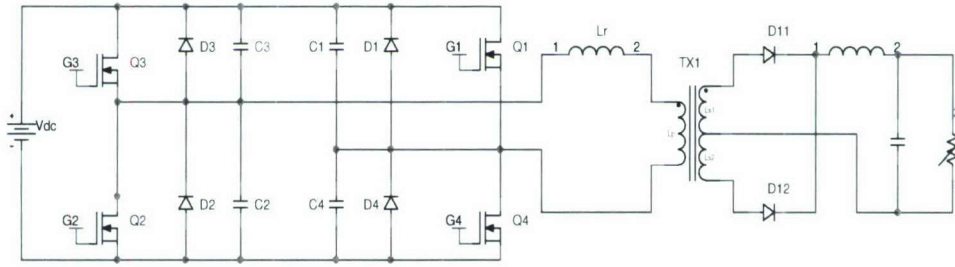


Fig 2. Load Current Dynamic Model

This model shows the relationship between the change in load current and the change in the output voltage. It is well-known that artificial neural networks can approximate any nonlinear function to the desired accuracy, after it is fully trained. In this research, the Simulink model is used to obtain data set for neural network training and testing. A feedforward neural network with six neurons in the hidden layer is employed. The activate function in the hidden layer is Tangent Sigmoid function:

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (6)$$

The weights of the neural network are initialized at random, and then updated by back-propagation algorithm. To speed up the training, Levenberg-Marquardt algorithm is employed:

$$W(k+1) = W(k) + \Delta W \quad (7)$$

where

$$\Delta W = (J_a^T J_a + \mu I)^{-1} J_a^T e \quad (8)$$

where  $J_a$  is the first order derivative of the error function respect to the neural network weight (also called the Jacobian matrix),  $e$  is the output error (i.e., the difference between the neural network output and the desired output),  $\mu$  is a learning parameter, and  $k$  is the number of iterations.

The neural network learning curve is shown in Fig. 3. From the simulation result we conclude that the neural network can successfully estimate the load current.

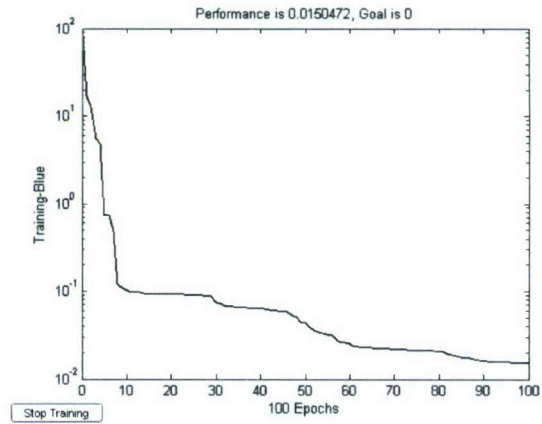


Fig 3. Training Performance on Load Current Estimation

#### IV. THE NEURAL NETWORK CONTROLLER DESIGN

Fig. 4 describes dynamic model for the overall system, with the estimation of load current, the input voltage, and the duty cycle. A feed-forward neural network with one hidden layer is employed to calculate the desired duty cycle. The inputs to the neural network include input voltage, load current, and the change of output voltage.

There are six neurons in the hidden layer and one neuron in the output layer. The activation function for each neuron is:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

The computer simulation results are shown in the following figures. Fig. 5 shows the training error is reduced to the desired accuracy after about 100 iterations. Fig. 6 demonstrates the duty cycle estimation of the neural network and the desired value. It is shown that once the neural network is trained, it is able to change the duty cycle based the change of the output voltage.

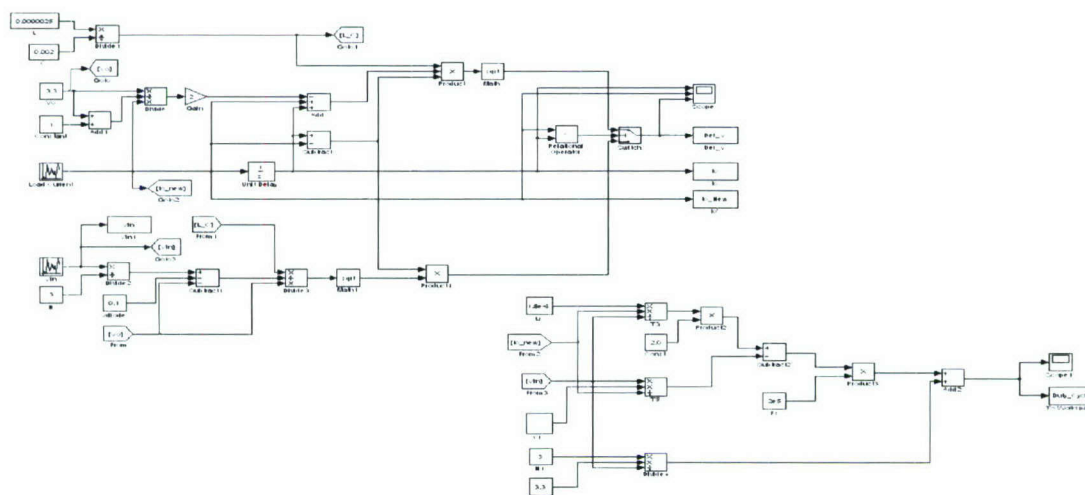


Fig. 4. The Simulink Model

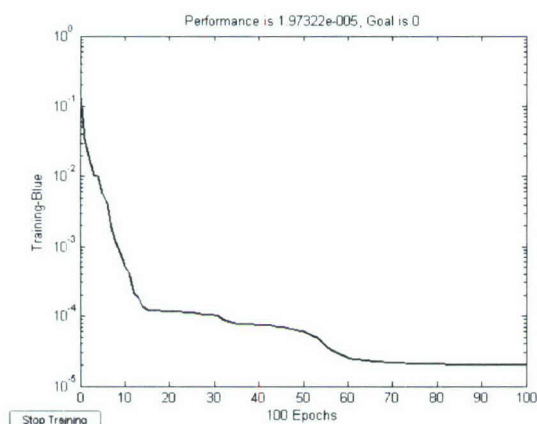


Fig. 5. Neural Network Training Error

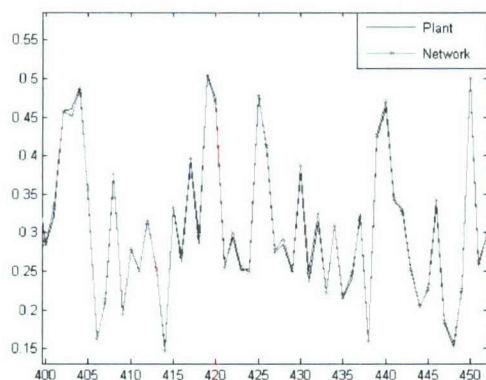


Fig. 6. Duty Cycle Estimation

## V. CONCLUSION

A neural controller as an alternative to classic controller for PSFB DC-DC converter is proposed in this paper. The simulation shows that the neural network controller is able to estimate the desired duty cycle under several dynamic conditions. More research works will be done to implement the neural network controller in hardware.

## ACKNOWLEDGEMENTS

This work was partially sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855.

## REFERENCES

- [1] P. R. Chetty, "Resonant power supplies: Their history and status," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 7, no. 4, pp. 23–29, Apr. 1992.
- [2] M. G. Kim and M. J. Youn, "An energy feedback control of series resonant converters," *IEEE Trans. Power Electron.*, vol. 6, no. 4, pp. 338–345, Jul. 1991.
- [3] J. M. Carrasco, E. Galván, G. E. Valderrama, R. Ortega, and A. Stankovic, "Analysis and experimentation of nonlinear adaptive controllers for the series resonant converter," *IEEE Trans. Power Electron.*, vol. 15, no. 3, pp. 536–544, May 2000.
- [4] H.-S. Choi and B. H. Cho, "Novel zero-current-switching (ZCS) PWM switch cell minimizing additional conduction loss," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 1, P. 165-172, February 2002.
- [5] H.-S. Choi, J.-W. Kim, and B. H. Cho, "Novel zero-voltage and zero-current-switching (ZVZCS) full-bridge PWM converter using

- coupled output inductor," *IEEE Trans. On Power Electronics*, Vol. 17, No. 5, P. 641-648, September 2002.
- [6] X. Ruan and Y. Yan, "A novel zero-voltage and zero-current-switching PWM full-bridge converter using two diodes in series with the lagging leg," *IEEE Trans. On Industrial Electronics*, Vol. 48, No. 4, P. 777-785, August 2001.
- [7] J. M. Quero, J. M. Carrasco, and L. G. Franquelo, "Implementation of a neural controller for the series resonant converter," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 3, P. 628-639, June 2002.
- [8] F. Kamran, R. G. Harley, B. Burton, T. G. Habetler, and M. A. Brooke, "A fast on-line neural-network training algorithm for a rectifier regulator," *IEEE Trans. On Power Electronics*, Vol. 13, No. 2, March 1998.
- [9] B. Choi, J. Kim, B. H. Cho, S. Choi, and C. M. Wildrick, "Designing control loop for DC-to-DC converters loaded with unknown AC dynamics," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 4, P. 925-932, August 2002.
- [10] Texas Instrument Inc, "BiCMOS advanced phase shift PWM controller," Data sheet of UCC3895, January 2001.

## **Visual speech analysis in real-life environment**

Project Investigator:

Jane (Xiaozheng) Zhang  
Electrical Engineering  
California Polytechnic State University  
San Luis Obispo, California

## **Abstract**

Automatic speech recognition technology has a variety of ongoing and potential applications in transcription and human-machine interaction. Mainstream automatic speech recognition (ASR) has focused almost exclusively on the acoustic signal. The performance of these systems degrades considerably in environments corrupted by acoustic noise and multiple talkers. One way to overcome this limitation is to supplement the acoustic speech with visual signal that remains unaffected in noisy environment. While previous research demonstrated that visual modality is a viable tool for identifying speech, the visual modality has yet to become utilized in mainstream ASR system. Most of the work to date has focused on visually clean data with uniform lighting and constant background. Hence there is a high demand for creating a robust visual front end in realistic environments. The primary goal of this research is therefore to develop statistical modeling of lip images in real life environment that allows a robust identification and tracking system that is flexible enough to handle variations in lighting, background, and human speakers. Emphasis is placed on real-world environment, thus models will be obtained and results will be evaluated using various databases with visually challenging data.

In our previous investigation, we have shown that the processing of color information is much faster than other features in identifying the lips. In addition, careful choice of color model can significantly improve efficiency and robustness of lip identification method. In particular, we have shown that hue color is a reliable feature because its color distribution is narrow and it has high discriminative power. We have performed statistical analysis of lip color models by examining hundreds of manually extracted lip images obtained from several databases. Our model thus represents lips taken by imaging devices with varying settings, in a variety of environments, and for a large group of population. In the next phase of our investigation, we will incorporate the developed lip model in classification of the lip pixels where several classic image filtering and segmentation techniques will be considered. In addition, several reliable lip features such as edge and shape will be included to increase the robustness of the lip identification algorithm.

## **Background**

Automatic speech recognition (ASR) allows a computer to identify the words a person speaks. For example, one can compose an email message, open a program using speech without touching the keyboard. A driver can select a radio station, operate an air conditioner by simply talking to the dashboard without distracting the eyes and hands. Automatic speech recognition can also be very helpful in assisting disabled individuals in their daily lives.

In the study of ASR, most mainstream research has focused exclusively on the acoustic signal. While purely acoustic-based ASR systems yield excellent results in a laboratory environment, the recognition error rate can increase dramatically in the real-world in the presence of noise, such as in a car environment with background noise, or in a typical office environment with ringing telephones and noise from fans and human conversations.

On the other hand, it is well known that humans with severe hearing losses use speechreading (lipreading) as the primary source of information for speech communication. It has been found that human listeners with normal hearing are also able to improve speech communication by seeing the speaker's face. Speechreading is particularly useful for noise-distorted speech. Motivated by humans' ability to speechread, research in automatic speechreading explores the role of visual speech information in the multimodal speech process and incorporates a visual component into an acoustic speech recognizer to improve recognition accuracy, especially in noisy conditions.

The first attempt to use vision to aid speech recognition was done by Petajan in 1984 [1]. He demonstrated that an audio-visual system outperforms either modality alone. During the past twenty years much progress was made in this area and various automatic speechreading systems were developed [2]. Those work clearly demonstrated that visual speech yields information that is not always present in the acoustic signal and enables improved recognition accuracy over purely acoustic-based systems, especially in environments corrupted by acoustic noise. Robust automatic speech recognition has long been an engineering goal for several decades. The use of the additional visual information has opened new possibilities.

However, the visual modality has not currently become incorporated in mainstream ASR system. The extra cost associated with processing video sequences is one practical concern. On the other hand, the robustness of the visual front-end design has yet to be justified in real-life environment. To-date, in almost all research in audio-visual ASR, work has concentrated on visually clean data – data collected in a studio-like environment. Those data were typically collected in the laboratory environment with controlled lighting conditions, uniform background and high video quality. The applicability of the developed visual front-end in a real-life environment bears close examination. For example, one of the most compelling places to embed speech is in the car. An “in-vehicle” voice response system promises safer driving through “hands-free, eyes-free” operation of the cell phone, dashboard controls and navigation system. Because of environmental conditions such as road noise and wind noise, acoustic-only speech recognition in a moving automobile is a very difficult problem. In this situation, supplementing the acoustic channel with visual speech information can have extreme potentials in improving speech recognition performance.

Visual speech recognition requires visual feature extraction. Tracking face and lips accurately in a moving automobile is challenging because of a variety of reasons. First the person might not be located ideally in the camera and the person's poses, head movement and person-camera distance might change from time to time. Second, the lighting condition within a car might change constantly. Extreme lighting and shadowing with changing background might cause the original visual front-end to fail. In addition, the video quality might not be of high resolution.

The primary goal of this research is therefore to develop statistical modeling of lip images in real life environment that allows robust identification and tracking systems.

Emphasis is placed on real-world environments, thus models will be obtained and results will be evaluated using various databases with visually challenging data.

### **Our Approach**

In our initial stage of the investigation, we seek to identify features of the lip image that are invariant to changing lighting conditions, background, and different speakers.

As a first step in this investigation, we examined the databases because the choice of a database in terms of its recording scenarios, variety and size of test subjects, imaging devices and camera settings, is of significant importance for the development and evaluation of the visual module. As mentioned above, one major limitation of the current automatic speechreading system is that it focused exclusively on data collected in studio-like environment. Currently, a number of corpora have been developed in the audio-visual speech research community: CMU database [3], XM2VTS [4], Tulips1 [5], DAVID [6], CUAVE [7], AVOZES [8], IBM database [9], BANCA [10], and AVICAR [11]. The first six databases are not suited for the purpose of this project because they are either gray-scale images, using lip highlighting, or are recorded in controlled environment. In CUAVE, different scenes can be added to the video backgrounds to simulate the effect of a person for example being in a moving car, however since the recording of the speakers took place in a controlled lighting environment, there is no lighting variation on a speaker's face. IBM database is a large audio-visual database covering 290 subjects in an office environment. However this database is a proprietary of IBM and they have not yet released the video data. Of all the databases, only the last two are of interest in this research. BANCA [10] is a large database with recordings of subjects in three different scenarios, controlled, degraded and adverse over a period of three months. Even though it was originally designed for training and testing of multi-modal verification systems, the data can be used for the purpose of lip-tracking. In AVICAR database [11], 100 speakers in an automobile environment were recorded under various driving conditions. This database represents a practical challenge for automatic speechreading system in a real-world condition. It is worth mentioning the Lombard phenomenon here. According to the Lombard effect [12], speech production is affected by the presence of the noise. People tend to speak differently when in a noisy environment – they speak louder, slower, more clearly, with more stress and emphasis in order to compensate for the noise [13]. However most of the current databases were recorded in an acoustically low noise environment and the various noise conditions are artificially created and added to video sequences later on. We believe that AVICAR database is mostly suited for this research objective because it was recorded under five different noise conditions, thus adequately addressed varying real-world test conditions. AVICAR was acquired in the project. We will purchase the BANCA in the next phase of our investigation. In addition to these two databases, we have also recorded our own data in office and home environments with varying lighting conditions including incandescent, fluorescent and daylight. For this purpose, we have acquired two cameras. One was an IREZ K2 web cam. The second camera was a Logitech Quickcam. We chose web cams over digital camcorder because they provide easy and quick interface to

computers. In addition, they are economical and still provide the resolution and functionality that we need to properly process images.

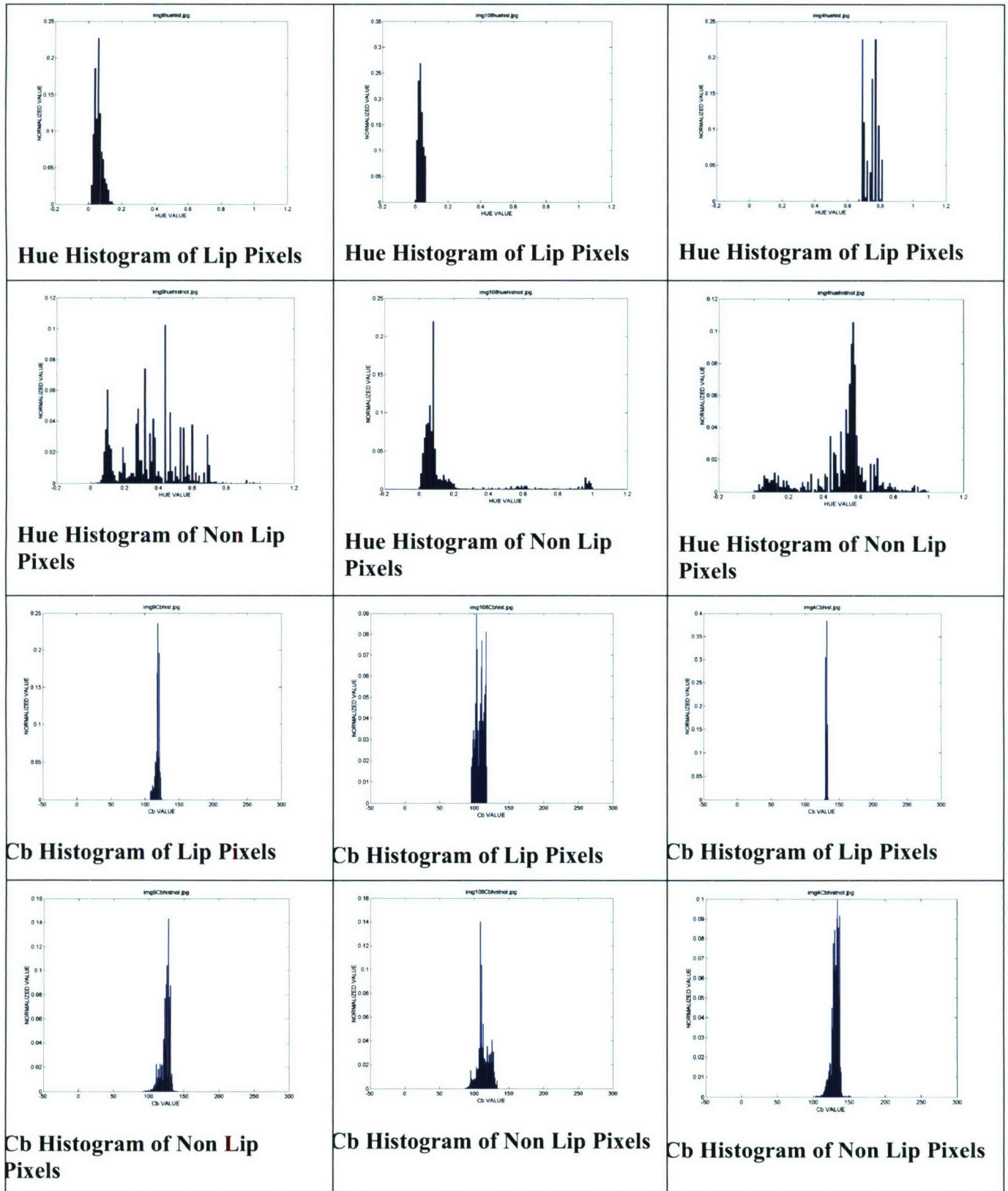
We have so far collected a total of 421 images in three different databases as part of our sample images. These images represent 129 different individuals of different cultures, genders and ages. Each image was taken under a fixed light type however the light intensity as well as the subjects position in the image varied. The three databases are divided as follows: 1.) Images collected using a Logitech web cam in a car or office environment (these images were collected by us); 2.) Images collected by browsing the Internet (these images were primarily face shots of celebrities so the images were most likely enhanced); 3.) Images collected by extracting still frames from the AVICAR [11] Database Videos. Figure 1 shows several characteristic images from each database.

In the second stage we performed a series of statistical analysis of lip features based on images from our databases. We identified color as an important feature for characterizing the object. Prominent colors can be used as a far more efficient search criterion for detecting and extracting the object, such as the red color for identifying the lips. In practice, a color image is represented in RGB color space. However, the RGB model is not able to separate the luminance and chromatic components of a color. This is a significant drawback since lips will then have different color values in environments with varying illumination. To circumvent the problem, we consider HSV space, which is a more intuitive way of specifying color that is based on human color perception. The HSV space is a non-linear transformation of the RGB space. In HSV color space each pixel has three components: hue, saturation, and value. Hue is the component that actually describes a pure color while value is the brightness component, which should absorb all illumination variations. We expect hue to have relatively good color constancy across varying conditions. Hence our first task in this project is to show whether hue can be used as a viable source for identifying the lip region.



**Figure 1: Characteristic Examples from Each Image Database**

We conducted statistical analysis on all available images to obtain a general understanding of how hue component of a lip image was affected by various environments. We were primarily concerned with the color data in the lip region of each

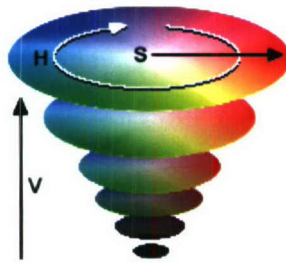


**Figure 2: Histograms for Images in Figure 1**

image as it compared to the background, therefore we began by manually cropping the lip region out of the image and creating histograms of each image - one containing only lip pixels and one containing all non-lip pixels. At this point we began to experiment with which color space would offer the greatest selectivity between the lip pixels and the non-lip pixels. For comparison, we consider both HSV and YCbCr. Not knowing ahead of time which of the two color spaces would offer the greatest level of selectivity, we constructed the corresponding histograms for both color spaces. Figure 2 shows histograms for the three images above.

Based on the histograms obtained from various sample images taken under various test conditions, we can make the following observations:

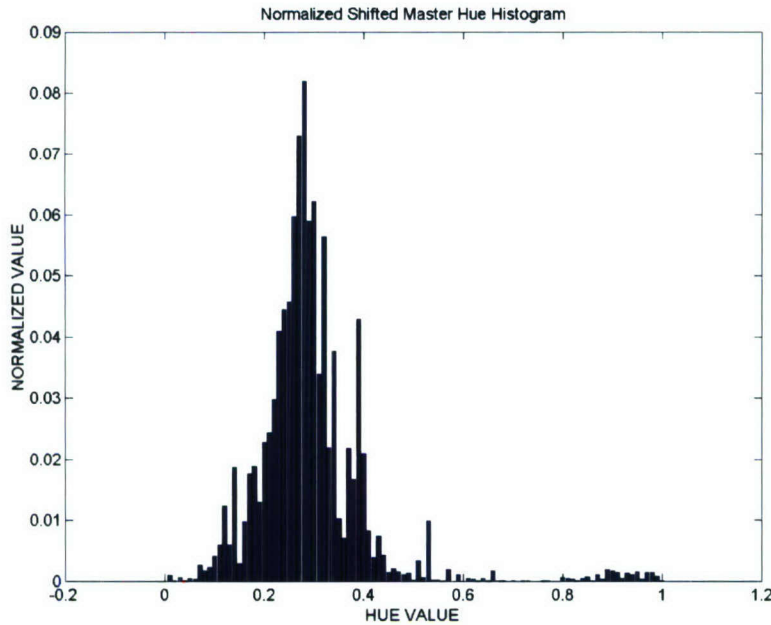
1. Hue has very good selectivity between lip pixels and background. In very few images did we find that the Hue distribution of the Lip region was entirely enveloped by the background. This means that in the HSV space it is easier to distinguish the lip pixels from the background. This is not true for the YCbCr space. As can be seen above, in almost all the images of our database the lip region histograms for both Cb and Cr values were entirely enveloped by the background. This would make it very difficult to distinguish a lip pixel from the background based only on color. Therefore the YCbCr color space was no longer considered as a useful color space for lip region segmentation purposes.



2. The Hue values of the lip regions were predominantly distributed in the color region we would expect to find red ( $0 < \text{Hue} < 0.2$  and  $0.8 < \text{Hue} < 1.0$ ). The reason that there are two continuous regions for the red Hue values is because of the nature of the HSV space to wrap around on itself in the red region.

3. The Hue histograms were invariant to the light intensity that was incident on the image subject. This though had some limitations. When the incident light was very intense as can be seen in the third image of Table 1, or if the incident light was very low, this tended to shift the histogram so that the lip pixels did not fall into the red region. This occurred in a minority of our images.

Finally the work was performed to acquire the mean Hue value and the Variance for the lip region of each image, and then find a Master Hue and Variance value that represents all images. In essence our theory is that the lip region can be described as a Normal Distribution in the Hue color space with a mean and variance that place the majority of lip pixels in the red region of the Hue color space. This can then be used to create a filter for segmenting the lip region from the background based solely on color. Figure 3 shows our Master Histogram. The Master Histogram was developed by creating one large hue vector, which contained the lip region hue values for every image in our database. From this vector we determined the Master Mean and Master Variance for our Gaussian Model.



**Figure 3: Master Histogram for Lip Hue**

### **Future work**

The future work of this project will focus further on lip region identification in varying lighting conditions. Now that we have created a working model for the hue distribution of human lips, we will create a segmentation process that is based on our Gaussian Model. This should be more effective than using a hard threshold based solely on the color value. In using this Gaussian Model we will employ several classic segmentation and image filtering techniques and include our application specific procedures.

### **References**

- [1] E. Petajan, *Automatic Lip Reading to Enhance Speech Recognition*. Ph.D. Thesis, University of Illinois, 1984.
- [2] G. Potamianos, et al. "Audio-Visual Automatic Speech Recognition: An Overview" in "Issues in Visual and Audio-Visual Speech Processing" by G. Bailly, E. Vatikiotis, and Perrier, Eds, MIT Press, 2004.
- [3] "URL: <http://amp.ece.cmu.edu/projects/audiovisualspeechprocessing>".
- [4] "URL: <http://www.ee.surrey.ac.uk/research/vssp/xm2vtsdb>".
- [5] J. R. Movellan, "Visual speech recognition with stochastic networks", In G. Tesauro, D.S.Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems*, Vol. 7, MIT Press, Cambridge, MA, 1995.

- [6] C.C. Chibelushi, S. Gandon, J.S.D. Mason, F. Deravi, and R.D. Johnston, "Design issues for a digital audio-visual integrated database," in *IEE Colloquium on Integrated Audio-Visual Processing for Recognition, Synthesis and Communication, Savoy Place, London, Nov. 1996*.
- [7] E. K. Patterson, S. Gurbuz, Z. Tufekci, and J.N. Gowdy, "CUAVE: A new audio-visual database for multimodal human-computer interface research", in *Proc. ICASSP*, 2002.
- [8] R. Goecke and B. Millar, "The Audio-Video Australian English Speech Data Corpus AVOZES," in *Proc. ICSLP 2004*.
- [9] G. Potamianos, E. Cosatto, H.P. Graf, and D. B. Roe, "Speaker independent audio-visual database for bi-modal ASR," in *Proc. European Tutorial and Research Workshop on Audio-Visual Speech Processing*, 1997.
- [10] E. Bailly-Baillire, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Boree, B. Ruiz, and J.P.Thiran, "The BANCA Database and Evaluation Protocol", in *Proc. IEEE ICASSP*, vol. 2, 2002.
- [11] Bowon Lee, Mark Hasegawa-Johnson, Camille Goudeseune, Suketu Kamdar, Sarah Borys, Ming Liu, Thomas Huang, "[AVICAR: Audio-Visual Speech Corpus in a Car Environment](#)", *INTERSPEECH2004-ICSLP*, October 2004.
- [12] E. Lombard, "Le signe de l'elevation de la voix", in *Ann. Maladiers Oreille, Larynx, Nez, Pharynx*, 37:101-119, 1911.
- [13] H. L. Lane and B. Tranel, "The Lombard sign and the role on hearing in speech", *Journal of Speech and Hearing Research*, (2):677-709, 1971.

### **Publications and Presentations Related to Project**

#### **Abstracts**

Zhang, Xiaozheng (Jane) and Montoya, Higinio Ariel; 'Statistical Modeling of Lip Color Features in Unconstrained Imagery'; *Presentation for 11<sup>th</sup> World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2007; July 8-11, 2007, Orlando, FL.*

Automatic Speech Recognition (ASR) allows a user to communicate with computers via voice instead of a mouse and keyboard. The use of speech interface in emerging multimedia applications is growing in popularity because it is more natural, easier, and safer to use. While ASR has witnessed significant progress in many well-defined applications, the performance of such systems degrades considerably in acoustically hostile environments such as in an automobile with background noise, or in a typical office environment with ringing telephones and noise from fans and human conversations. One way to overcome this limitation is to supplement the acoustic speech with a visual signal that remains unaffected in a noisy environment. While previous research demonstrated that visual modality is a viable tool for identifying speech, the

visual modality has yet to become utilized in mainstream ASR systems. One obstacle is the difficulty in building a robust visual front end that tracks lips accurately in a real-world condition. To date majority of the work in automatic speech reading has focused on databases collected in studio-like environments with uniform lighting and constant background. Accurately detecting and tracking lips under varying environmental conditions and for a large group of population is a very difficult task due to large variations in illumination conditions, background, camera settings, facial structural components (beards, moustaches), and inherent differences due to age, gender, and race. In this paper we present our current progress in obtaining a robust visual front end in real-world conditions. We examine the use of color information in detecting the lip region and derive the statistical modeling of lip images by examining hundreds of manually extracted lip images obtained from several databases. We also report our preliminary results on the lip image segmentation using the obtained models.